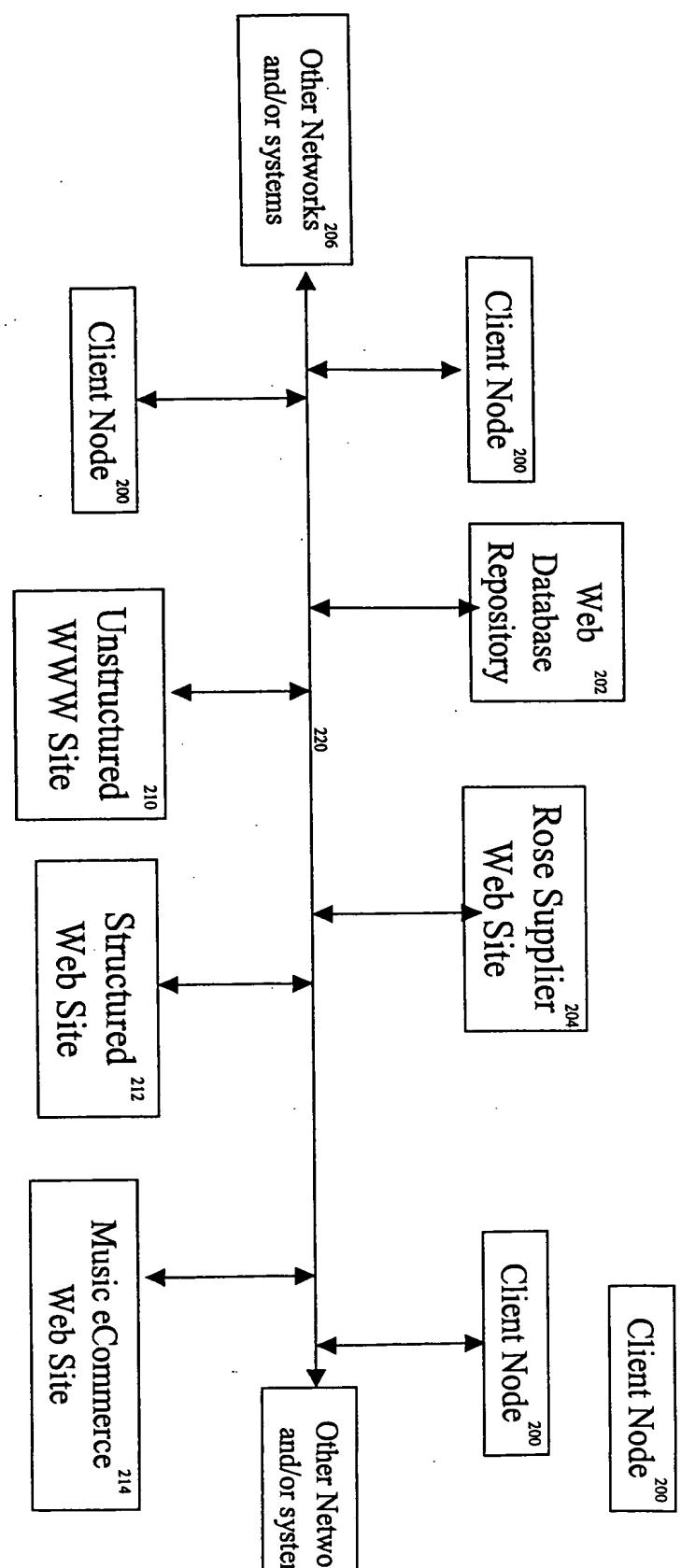


Computer Environment
Figure 1



Shows a typical embodiment where the invention is installed into a plurality of Client Nodes extracting data from a plurality of data repositories.

Figure 2

Example WWW Page

Next County ³⁰⁰
Previous State ³⁰²

This text is of no interest to us. It could be contained in a separate table, frame or other HTML container such that we are able to identify its boundaries (ie the start and end of the text field) and such that we can determine that it is not anything of use.

\$55,000 for a great 3 bedroom 2 bath house. MLS1721

³⁰⁴

Banner Advertisement

both content and URL link changes between page accesses.

³⁰⁶

Mortgage ³⁰⁸
Calculator

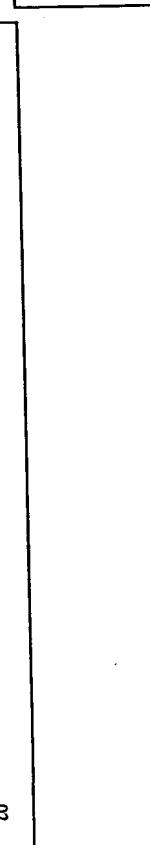
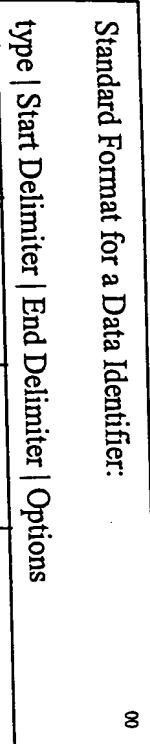
Regional Information representing data of *no interest* as we are looking for Home price and descriptions and not information on the area.

³⁰⁸

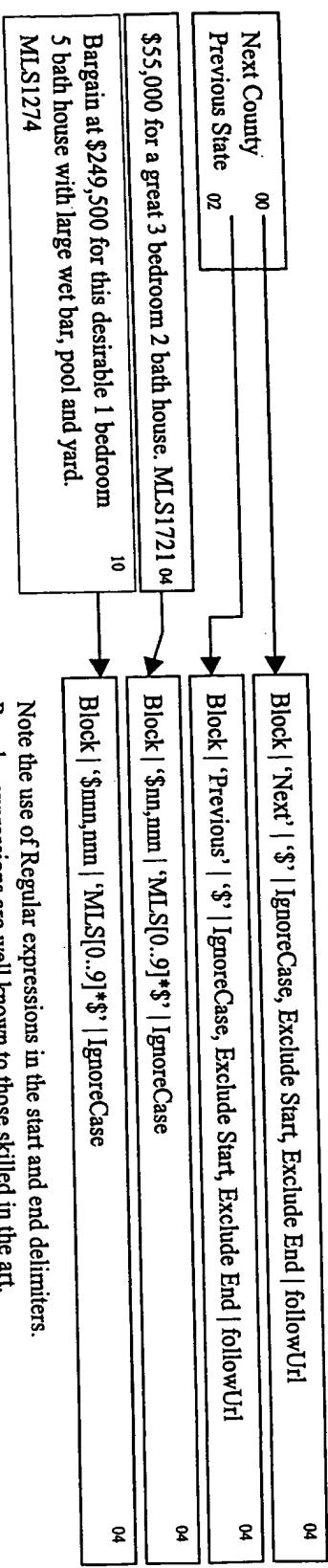
Bargain at \$249,500 for this desirable 1 bedroom 5 bath house with ³¹⁰
large wet bar, pool and yard. MLS1274

Example WWW page with data of *interest* and data of *no interest*.

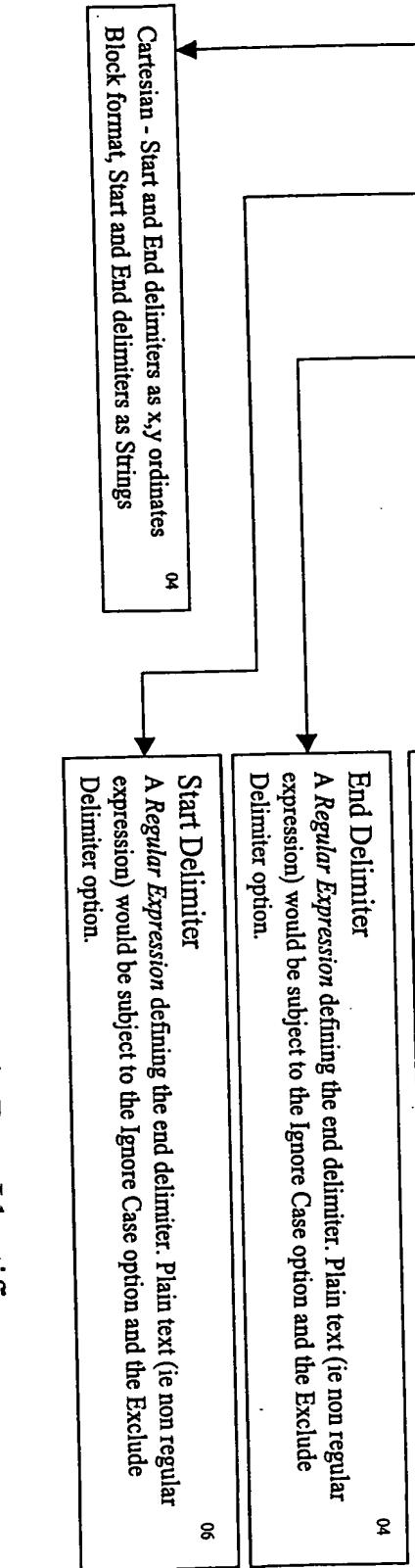
Example WWW page
Figure 3



Text to Decode

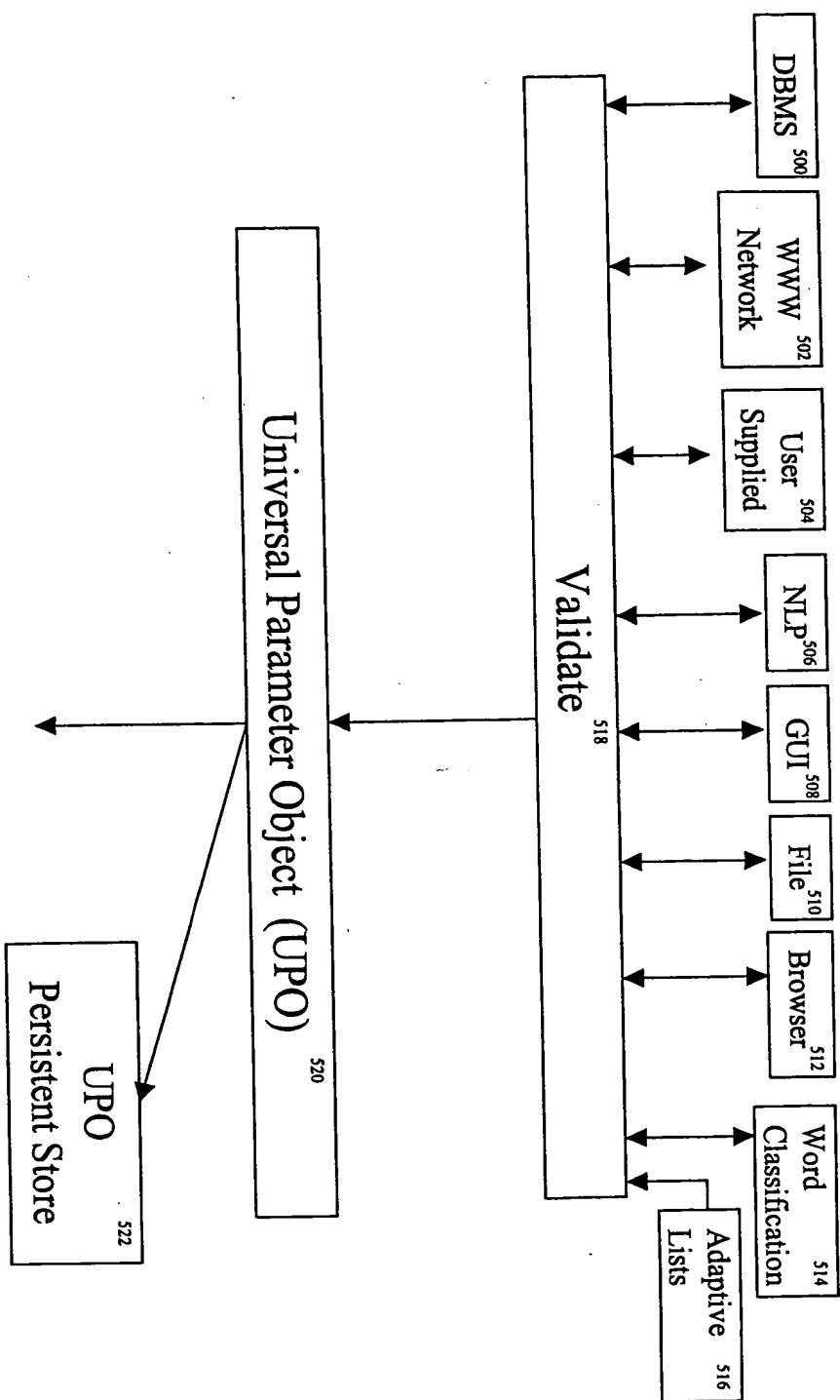


Example Data Identifier

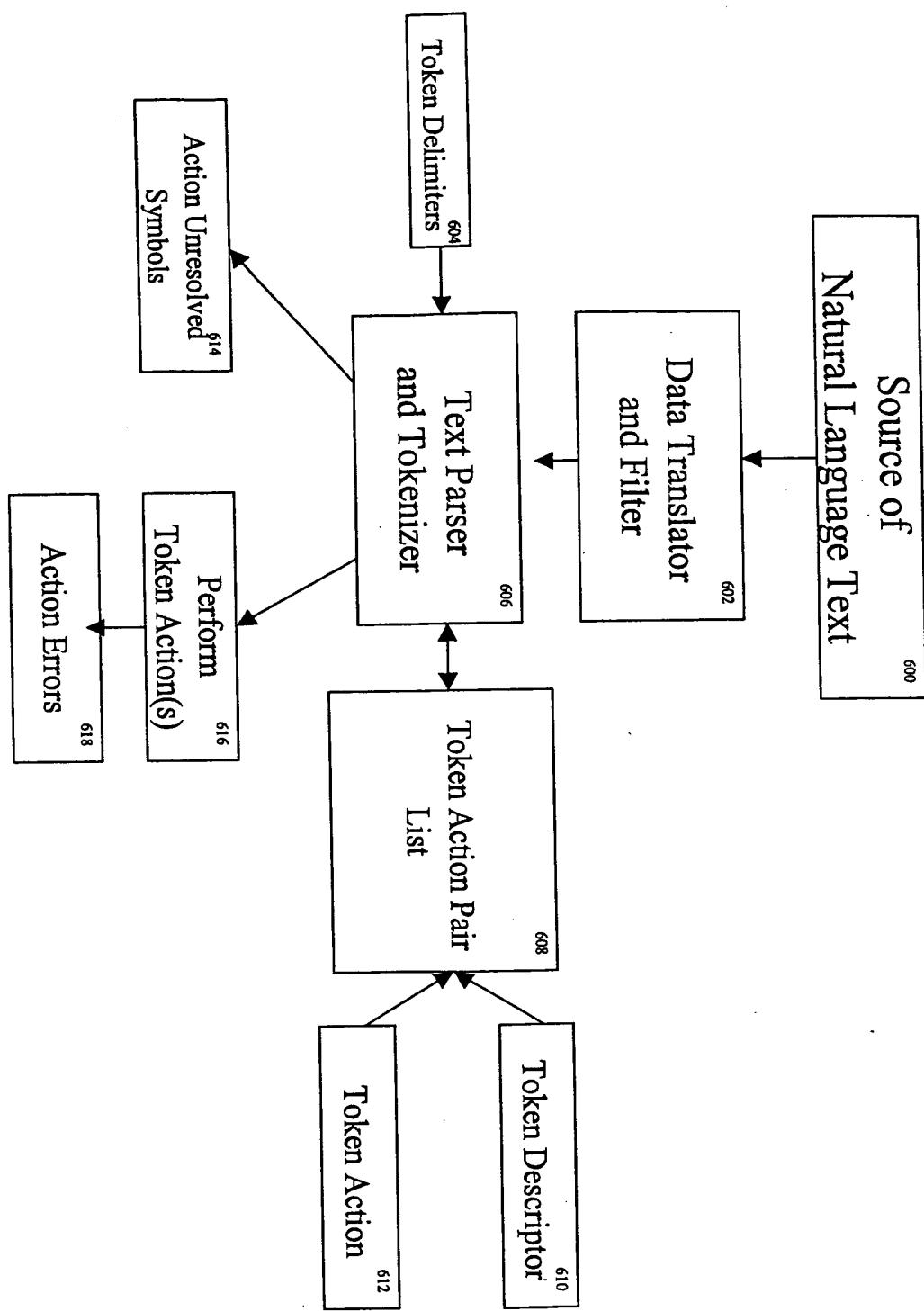


Example Data Identifier Usage
Figure 4

Universal Parameter Object (UPO)

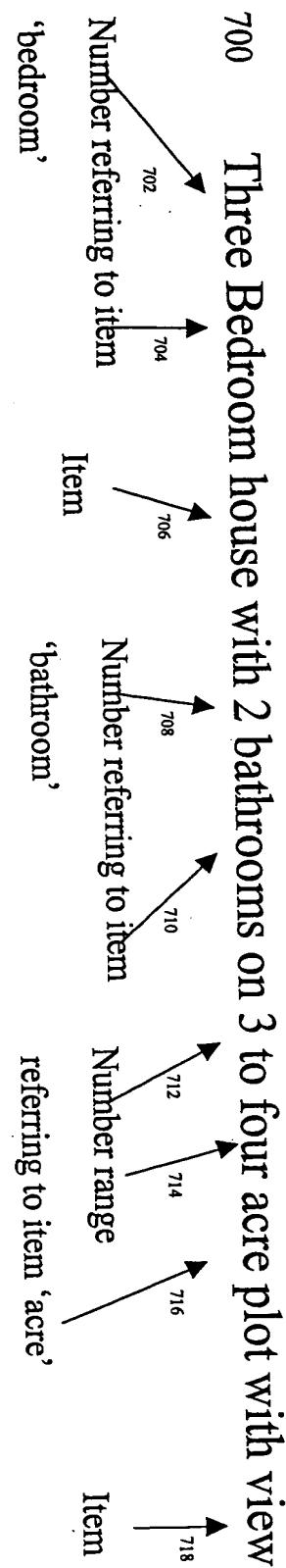


Universal Parameter Object (UPO)
Figure 5

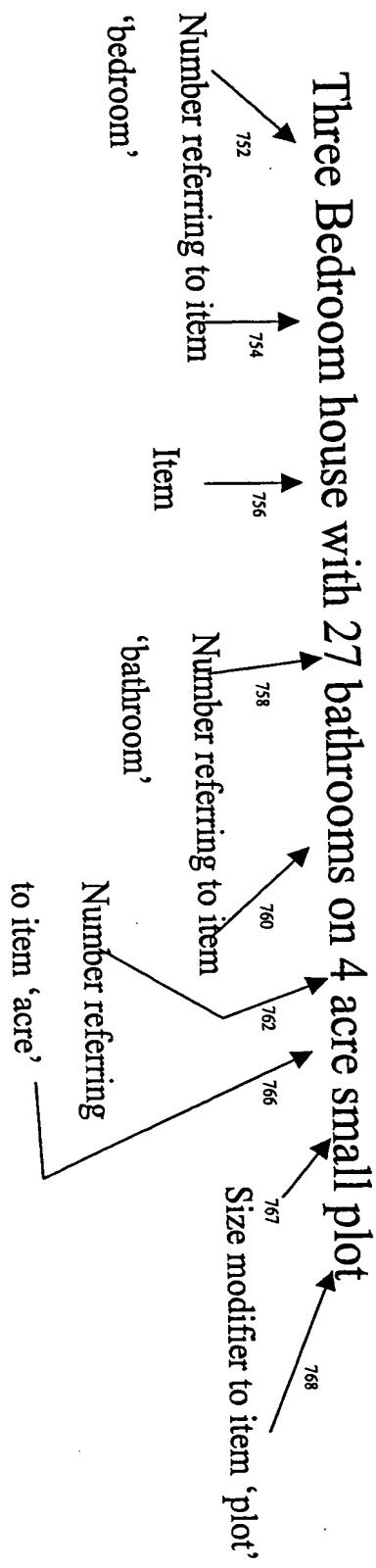


Natural Language Processor
Figure 6

An example of Natural Language with parameters:-



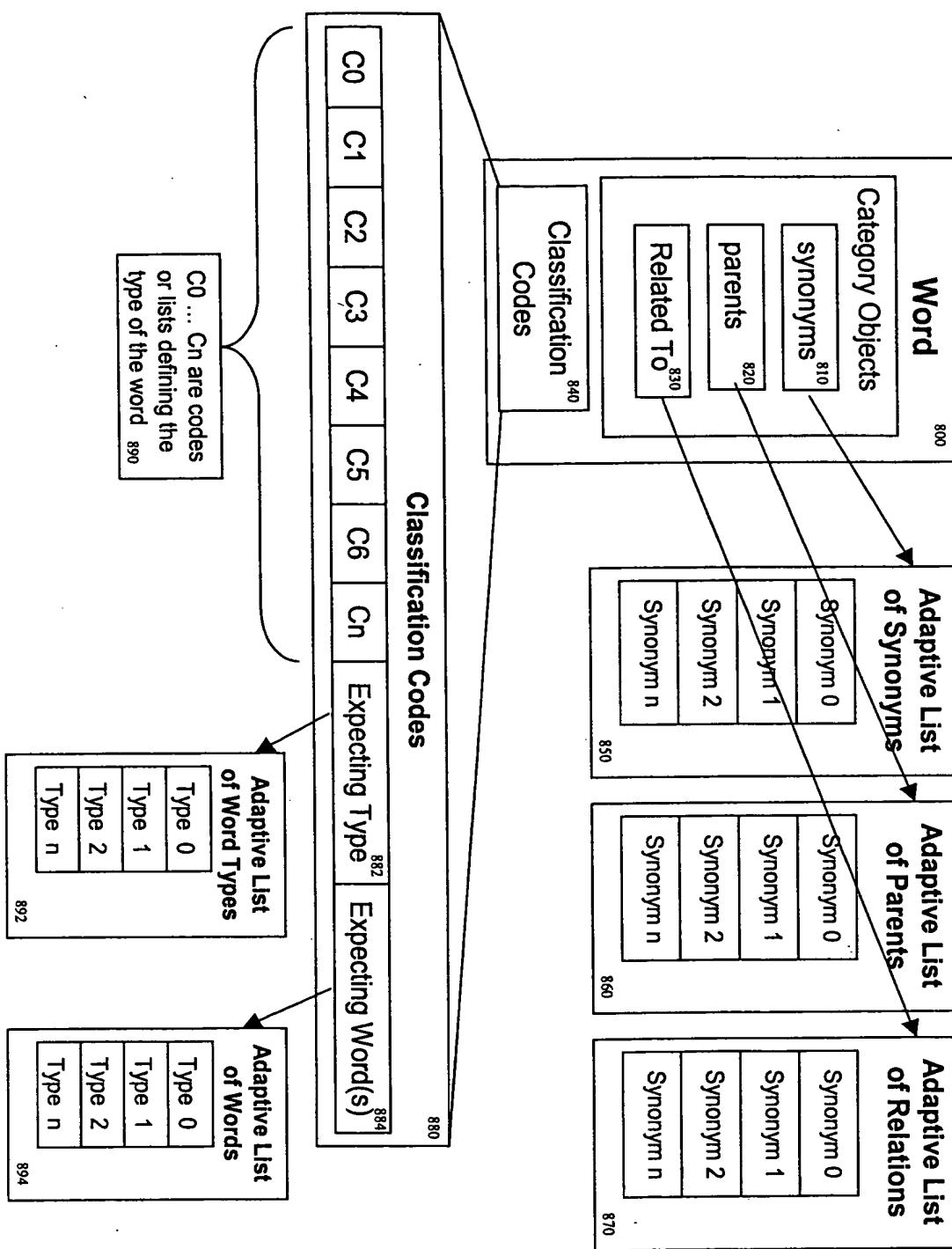
750 An example of Natural Language with improbable and conflicting parameters:-



Examples of Natural Language

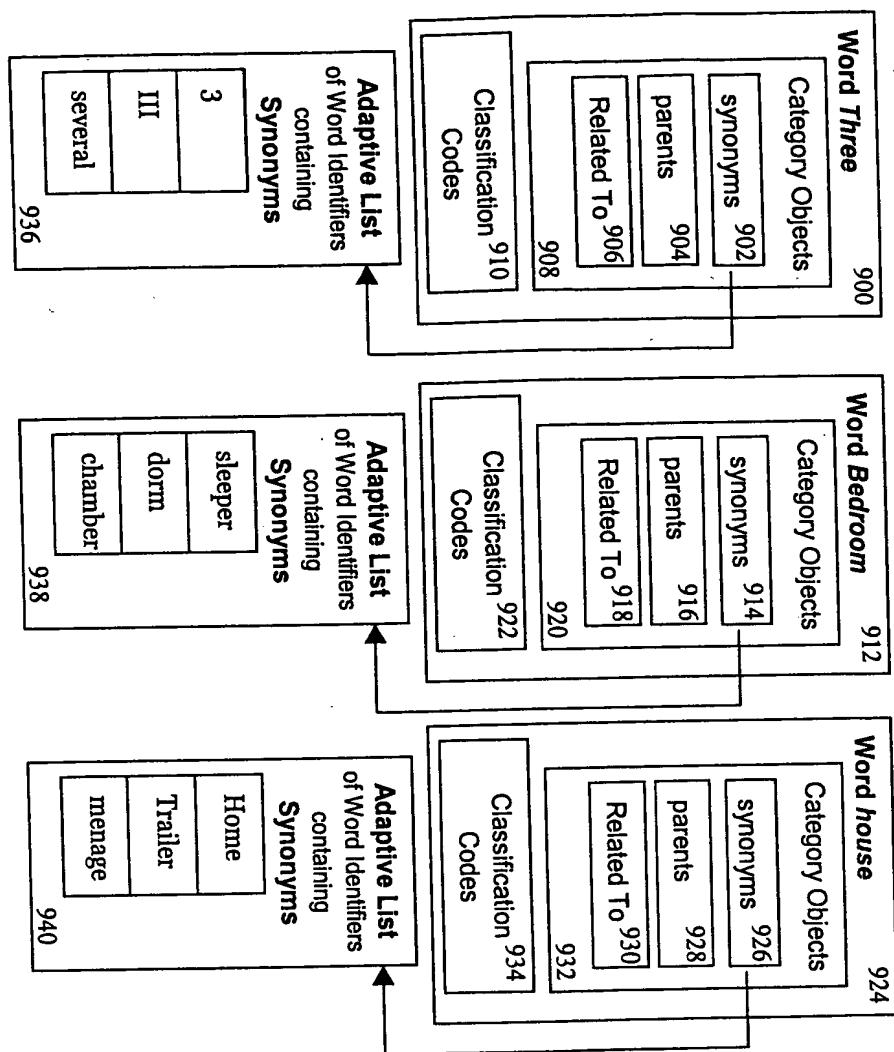
Figure 7

Word Identifier

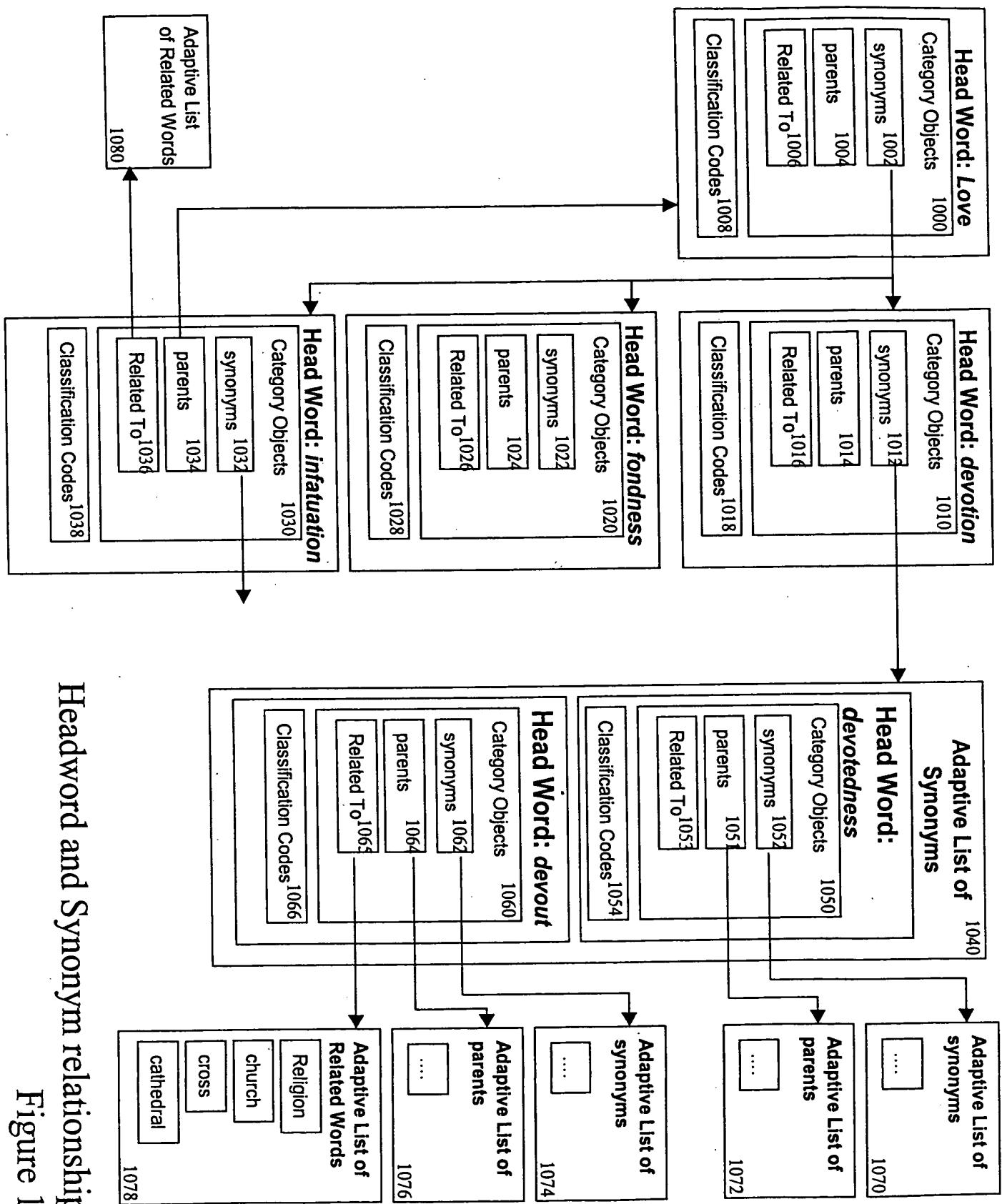


Word Identifier
Figure 8

Three Bedroom house with 2 bathrooms on 3 to four acre plot with view



Natural Language to Word Ident Mapping
Figure 9



Headword and Synonym relationships
Figure 10

Word Classification

Category Codes							
C0	C1	C2	C3	C4	C5	C6	Cn
							Expecting Word(s)

1100

Target Word: "Cat"

Example word "cat" showing example categories it fits into. Such categories will vary between embodiments.

Natural Category 1002 1110	Gender 1 1112	Species 1223 1114	Food Group 1000 1116	Function 5002 1118	Specific 0 1120	Specific n 1122
----------------------------------	---------------------	-------------------------	----------------------------	--------------------------	--------------------	--------------------

Example Natural Category Codes Flaura (2000) Fauna (8000) Mammal (1002) Insect (1003) Reptile (1004)	Example Gender Codes None 0 Female 1 Male 2	Example Species Codes Felis catus (1223)	Example Food Group Codes Carnivore (1000) Herbivore (2002)	Example Function Codes Sleeps (5002) Worker (6000)
---	--	---	--	--

The number, meaning and definition of categories will be dependant on the specific embodiment.

Word Comparison

Category Codes

C0	C1	C2	C3	C4	C5	C6	Cn	Expecting Type	Expecting Word(s)
----	----	----	----	----	----	----	----	----------------	-------------------

120

Item 1: "Cat"	Natural Category 1002	Gender 0	Species 1223	Food Group 1000	Function 5002	1212	Comparison of the codes for the words "cat" and "tulip". The difference is large as obviously a "cat" is an animal and a "tulip" is a flower.
Item 1 : "Tulip"	Natural Category 8000	Gender 0	Species 7014	Food Group 3007	Function 9668	1214	
Difference (distant match)	7000	0	5791	2007	4666	1216	

1210

Item 1: "Cat"	Natural Category 1002	Gender 0	Species 1223	Food Group 1000	Function 5002	1220	Comparison of the codes for the words "cat" and "lion". The difference is very small as obviously a "cat" is closely related to a "lion"
Item 1 : "Lion"	Natural Category 1002	Gender 0	Species 1227	Food Group 1000	Function 5003	1222	
Difference (Almost exact match)	0	0	4	0	1	1224	

1226

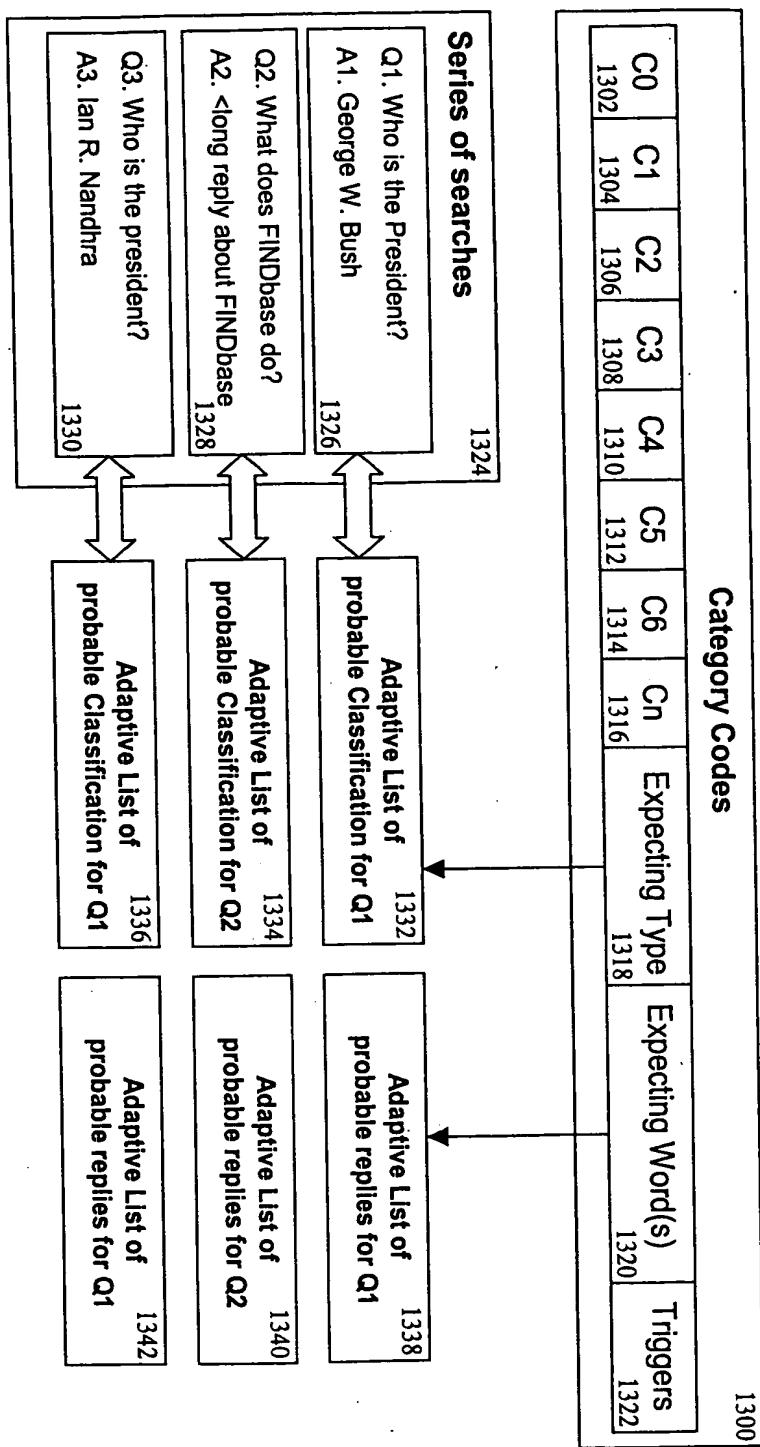
$$1230 \text{ Difference } \Delta_{c0} = \frac{1}{c0} - \frac{1}{c1}$$

$$\text{Proximity } T = \frac{1}{c0} - \frac{1}{c1}$$

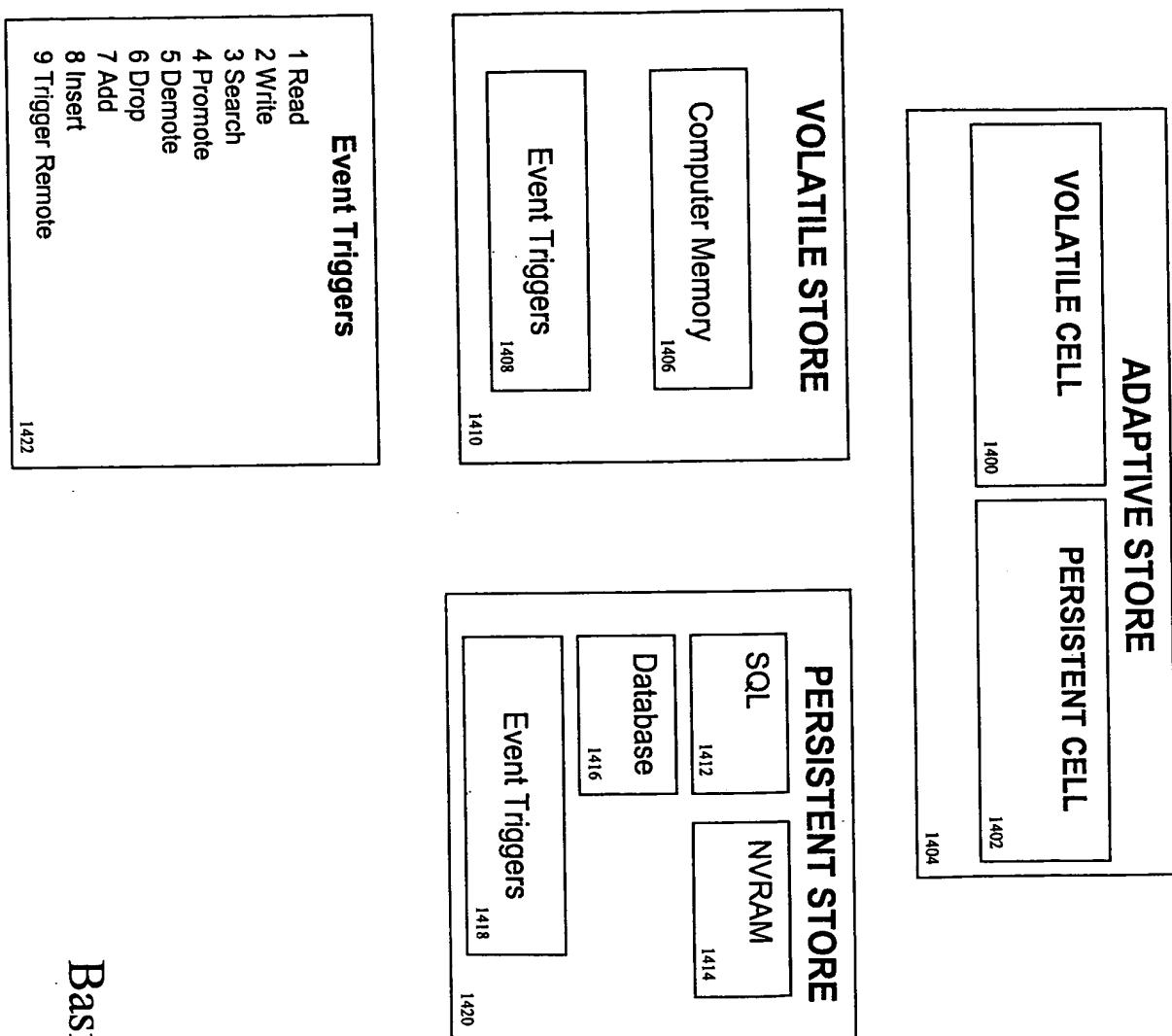
1236

1232 Difference $\Delta_{c1} = \frac{0}{c1} - \frac{1}{c1}$	$\frac{cx}{cn} \Delta_{cx}$	Where cx and cn encompass a set or sequence of cells such as c0, c3, c4, c5, c9	1238
1234 Difference $\Delta_{cn} = \frac{0}{cn} - \frac{1}{cn}$			

Word Expectation(s)

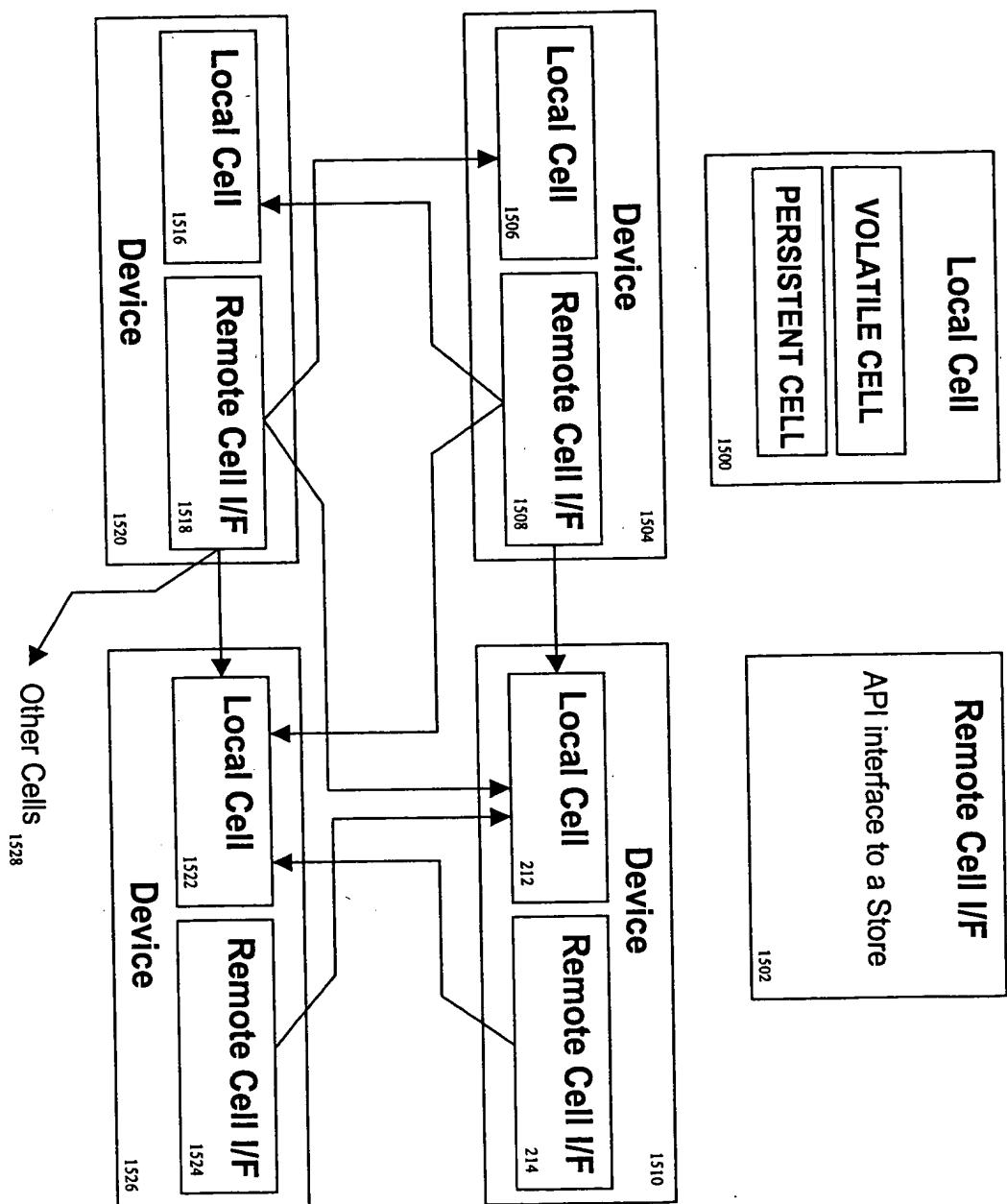


Basic Stores



Basic Store Types
Figure 14

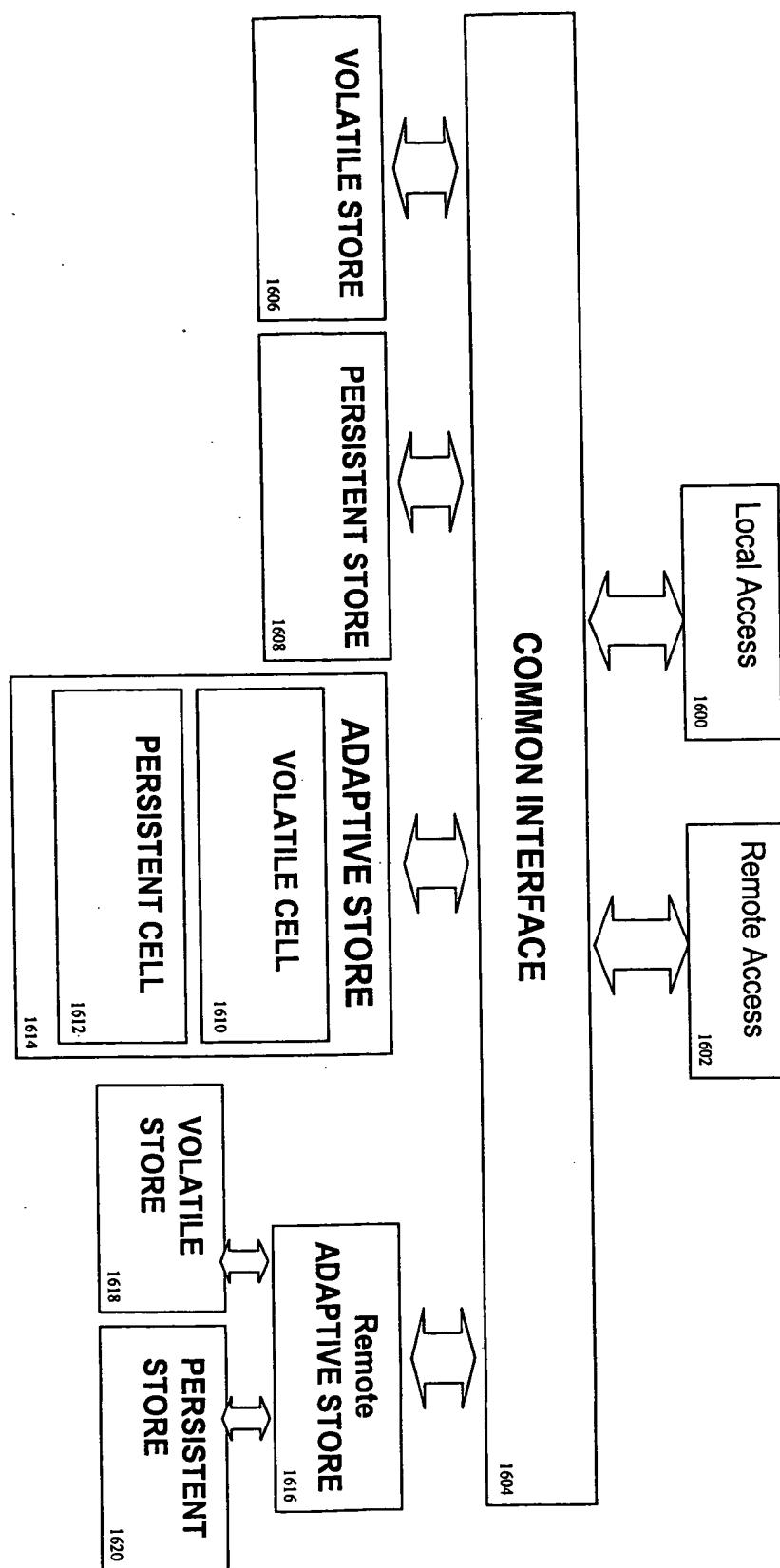
Storage Cells



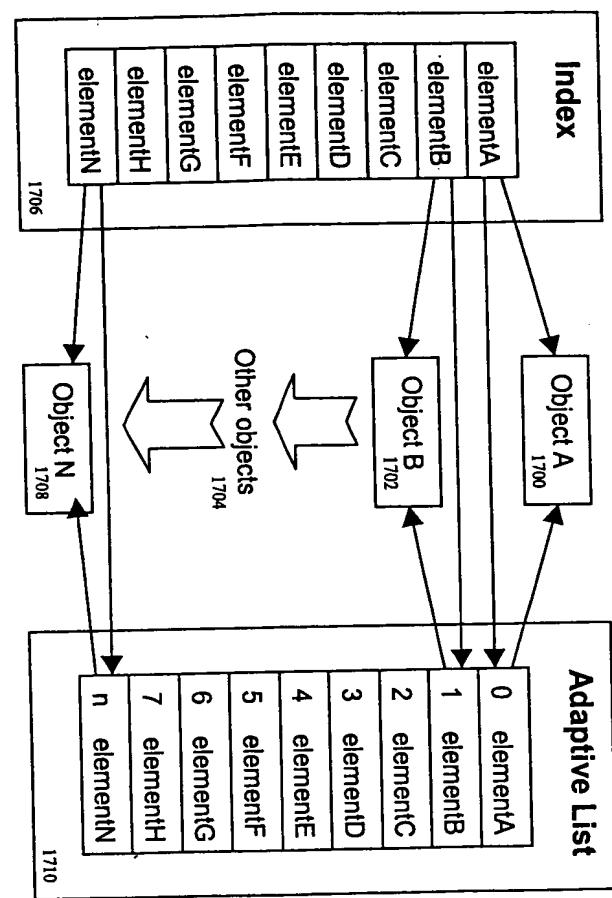
Storage Cells
Figure 15

Storage API

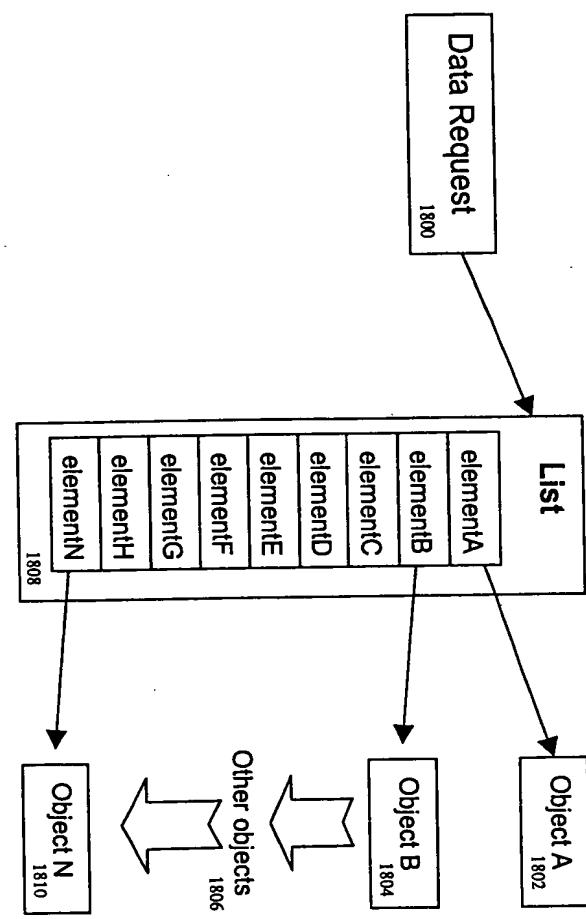
10/517738



Adaptive Store



Adaptive Store Indexing and List relationship
Figure 17



Adaptive Store – Bare Storage
Figure 18

Adaptive Store – Simple Accesses

Initial State of Adaptive List
 After 1st Search for element D
 After 2nd Search for element D
 After 3rd Search for element D
 After Search for element G
 After Search for element H

Adaptive List

0	elementA
1	elementB
2	elementC
3	elementD
4	elementE
5	elementF
6	elementG
7	elementH
n	elementN

0	elementA
1	elementD
2	elementB
3	elementC
4	elementE
5	elementF
6	elementG
7	elementH
n	elementN

0	elementA
1	elementD
2	elementB
3	elementC
4	elementE
5	elementF
6	elementG
7	elementH
n	elementN

0	elementD
1	elementA
2	elementB
3	elementC
4	elementE
5	elementG
6	elementH
7	elementF
n	elementN

0	elementD
1	elementA
2	elementB
3	elementC
4	elementE
5	elementG
6	elementH
7	elementF
n	elementN

After addition of new element Z

Least accessed element N dropped and replaced by new element Z

Accessed elements are elevated one level in the list and the element above is demoted by one element.

19/39

Adaptive Store Simple Access
 Figure 19

0	elementD
1	elementA
2	elementB
3	elementC
4	elementE
5	elementG
6	elementH
7	elementF
n	elementZ

1912

After addition of new element Z

Least accessed element N dropped and replaced by new element Z

Accessed elements are elevated one level in the list and the element above is demoted by one element.

1914

1916

Adaptive Store – Weighted Accesses

Initial State of Adaptive List After 1st Search for element E After 2nd Search for element E After 3rd Search for element E After Search for element B

Adaptive List		Adaptive List		Adaptive List		Adaptive List	
0	Priority 4 elementA	0	Priority 4 elementA	0	Priority 4 elementA	0	Priority 5 elementB
1	Priority 4 elementB	1	Priority 4 elementB	1	Priority 4 elementB	1	Priority 4 elementA
2	Priority 40 elementC	2	Priority 40 elementC	2	Priority 41 elementE	2	Priority 41 elementE
3	Priority 4 elementD	3	Priority 39 elementE	3	Priority 40 elementC	3	Priority 40 elementC
4	Priority 38 elementE	4	Priority 4 elementD	4	Priority 4 elementD	4	Priority 4 elementC
5	Priority 4 elementF						
n	Priority 0 elementN						

2000

2002

2004

2006

2008

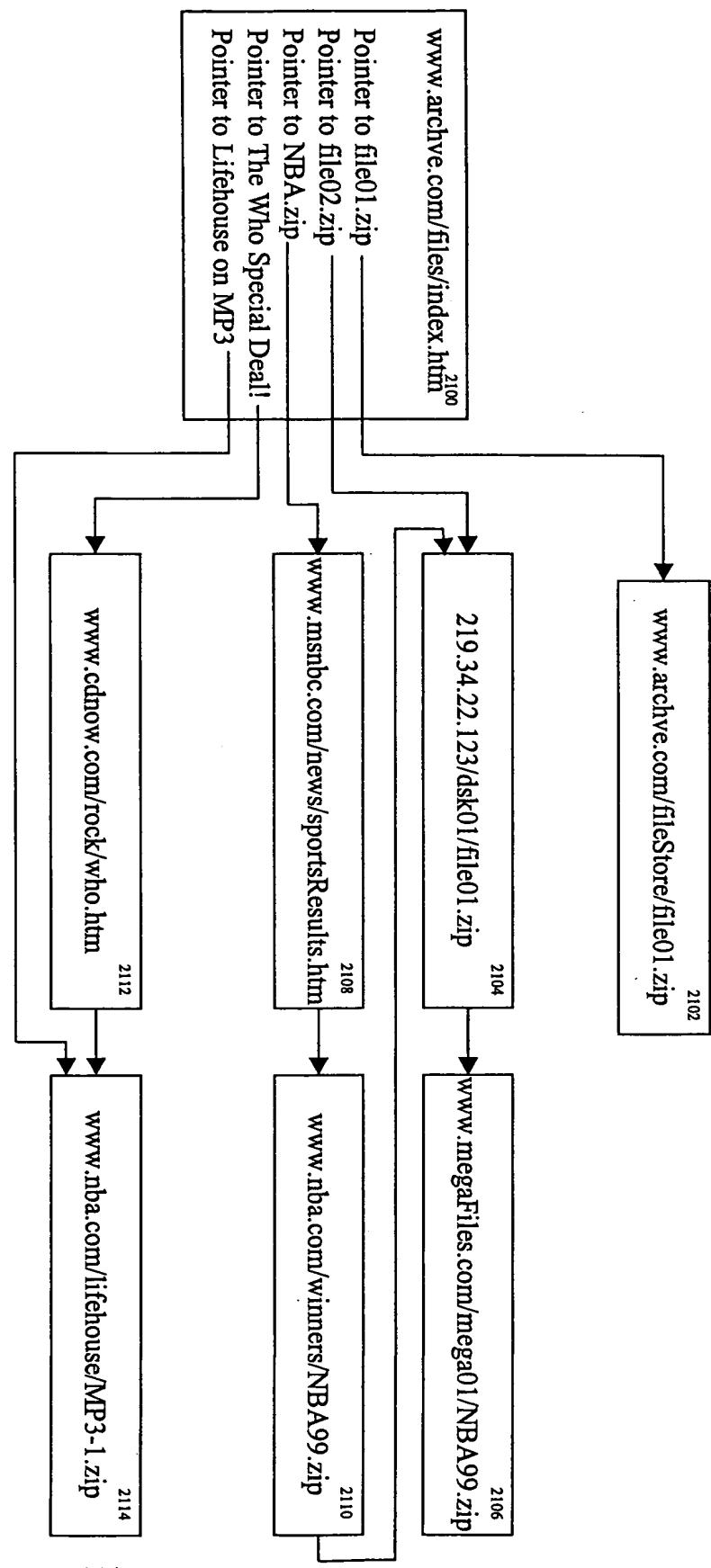
After addition of new element Z 2022

Accessed elements have their weight values incremented. If the new priority value is greater than the element at location n-1, element n is swapped with element n-1
2024

Least accessed element N dropped and replaced by new element Z 2026

Adaptive Store Weighted Access

Figure 20



Examples of how links to files can span different repositories on a Network such as the WWW and how such links may reference the same file.

URL Chaining and Indirection
Figure 21

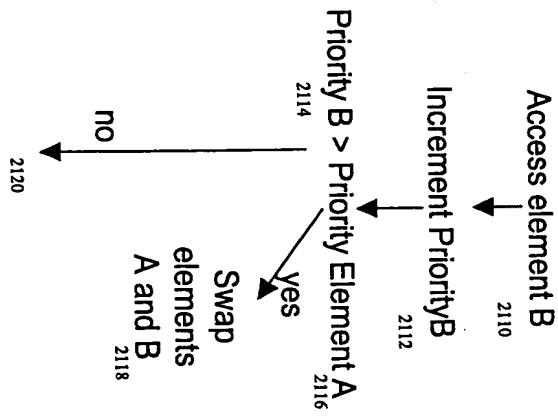
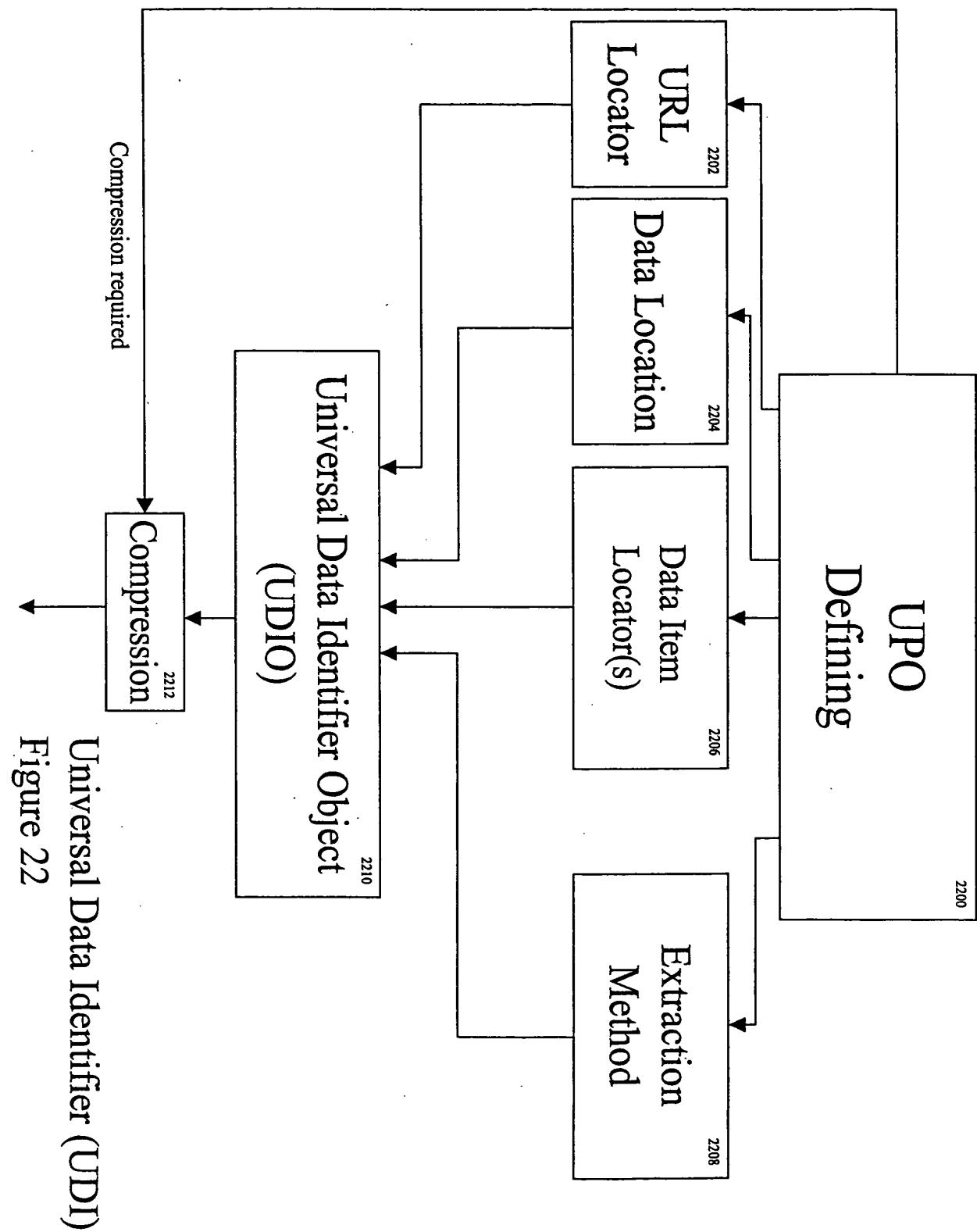
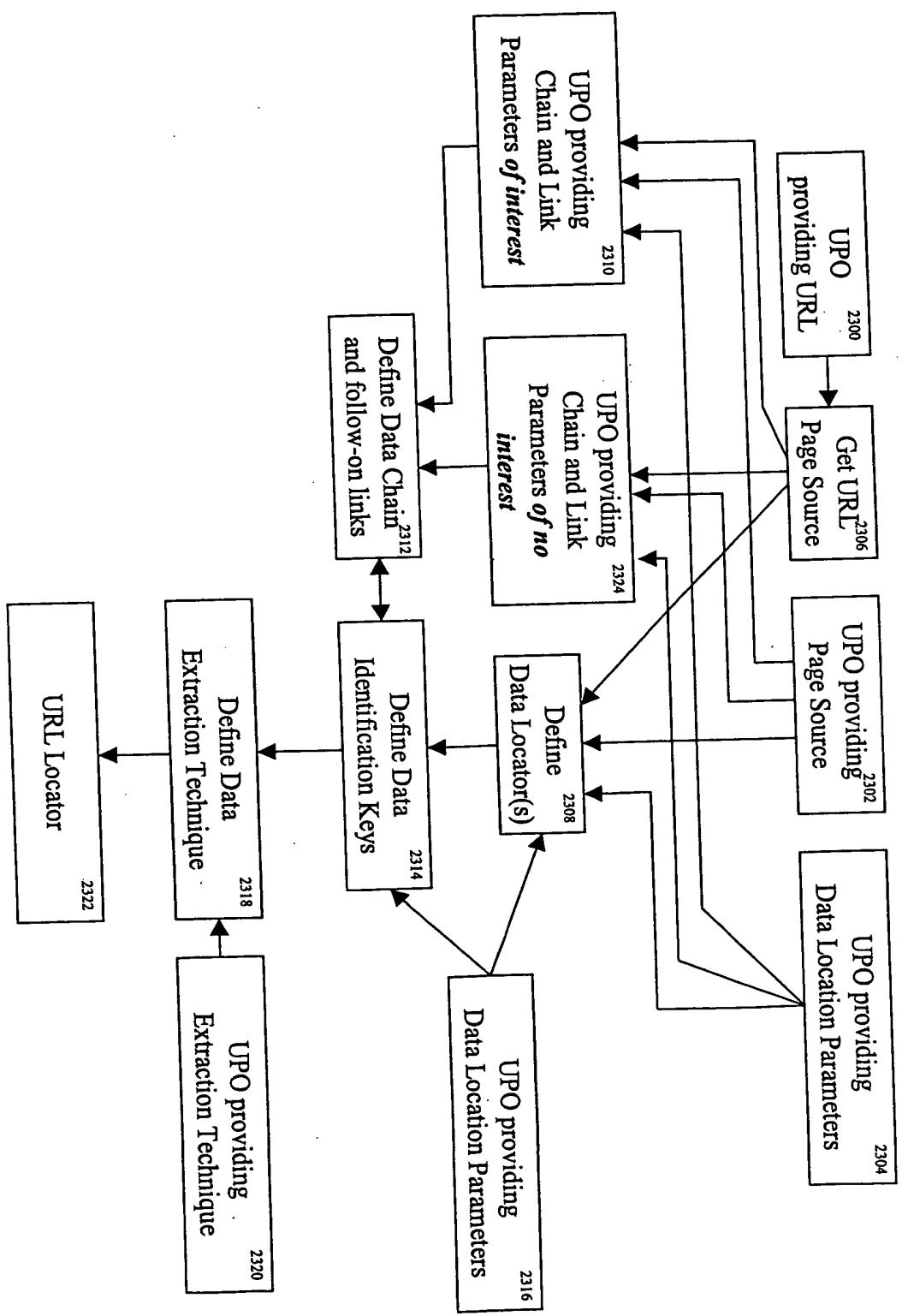


Figure 20-1

Universal Data Identifier (UDI)

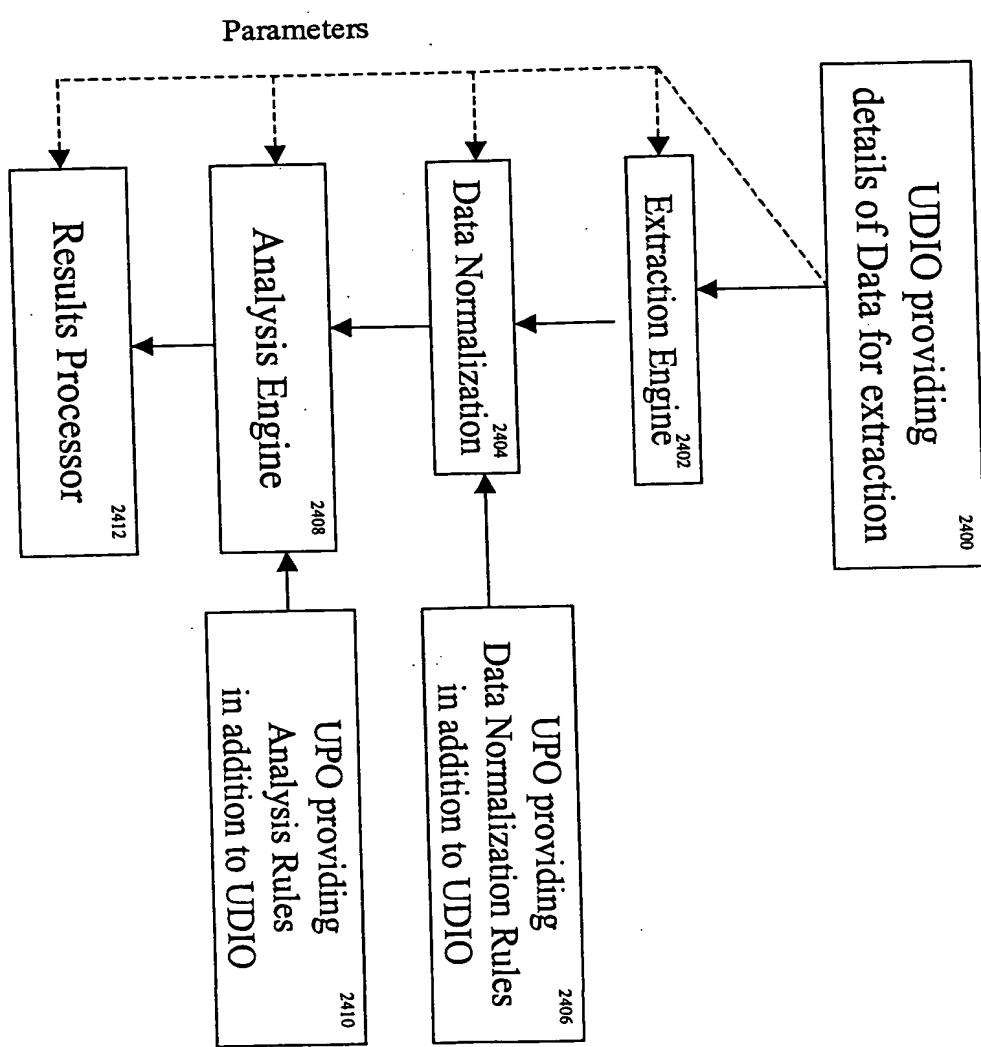


Universal Data Identifier (UDI)
Figure 22

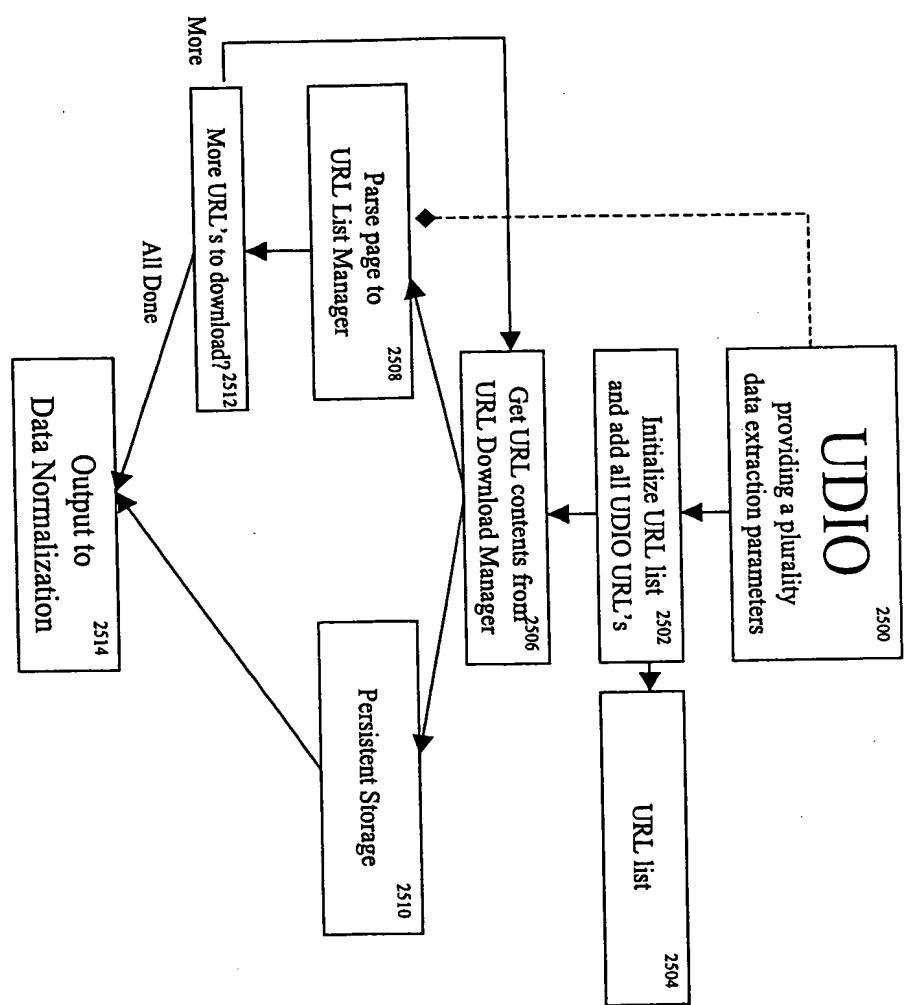


URL Universal Data Identifier

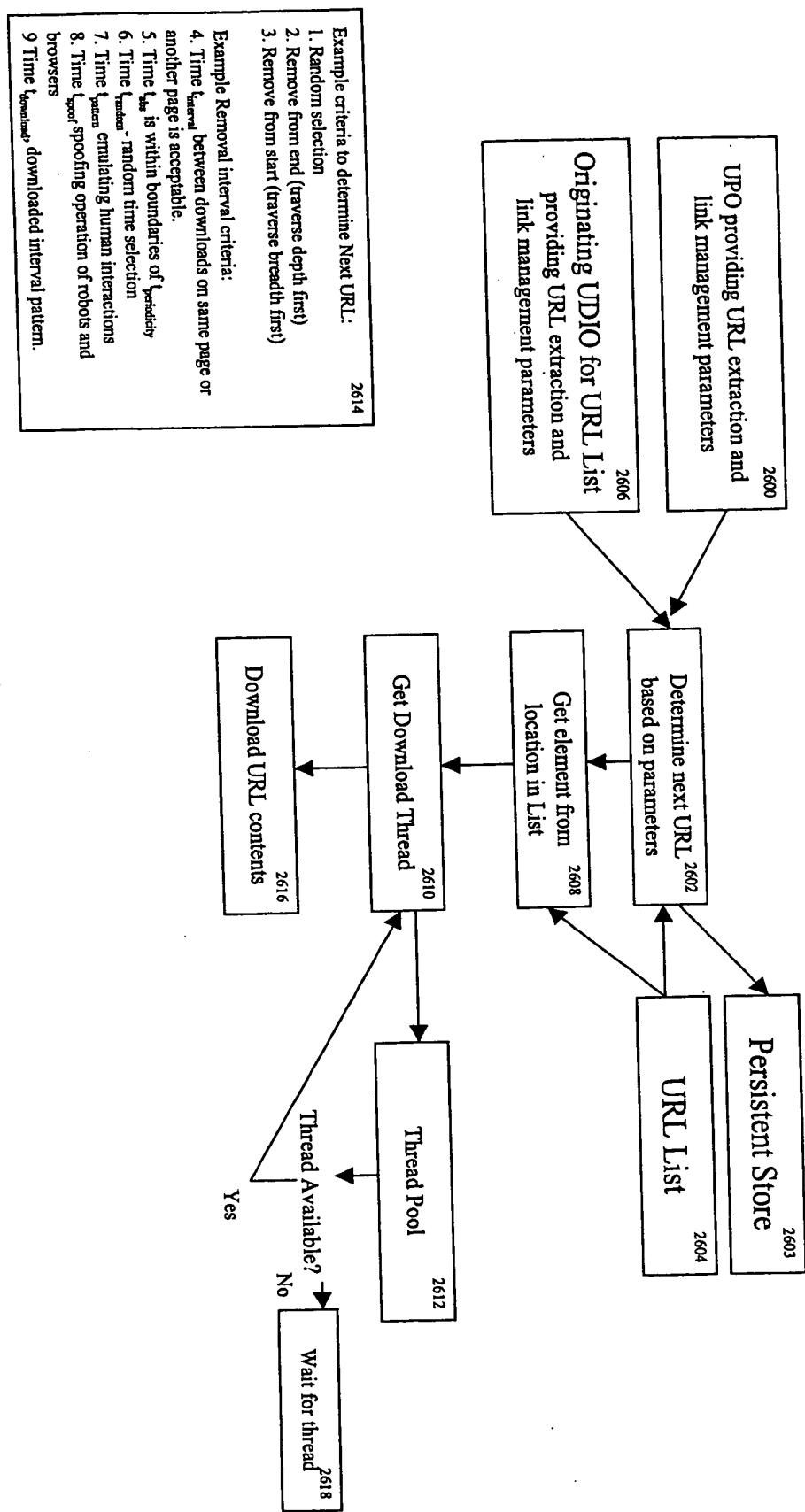
Figure 23



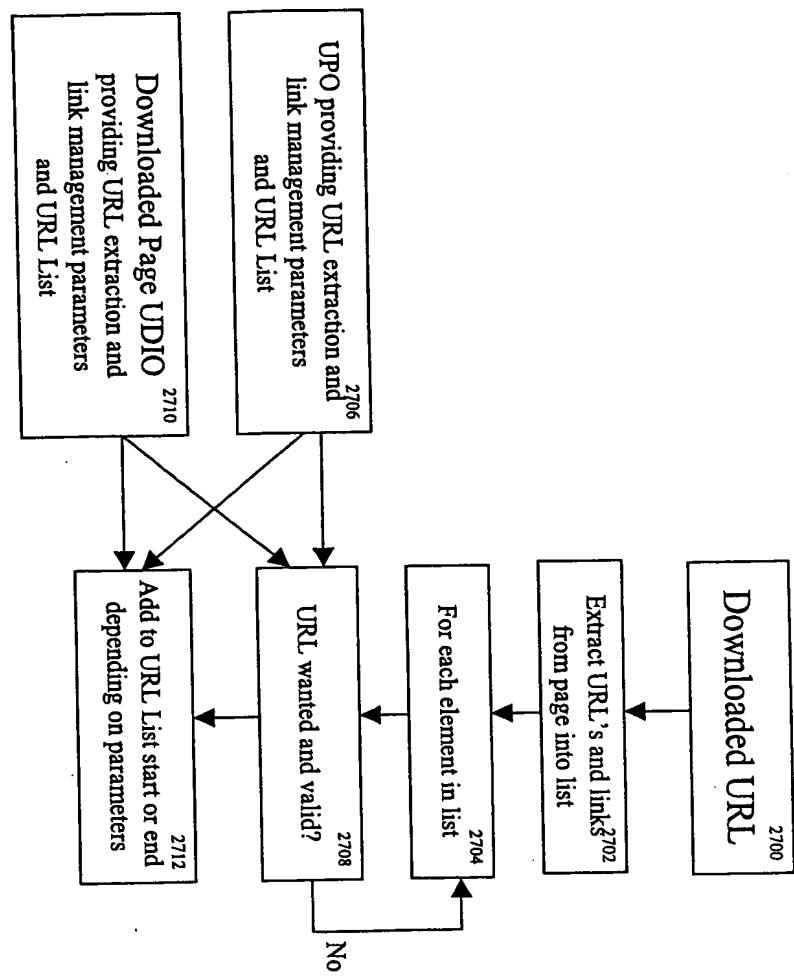
Client Data Extractor (CDE)
Figure 24



WWW Extraction Engine,
Figure 25



URL Download Manager
Figure 26



URL List Manager
Figure 27

Analyzed Data

Format Switch

Criteria Selector UPO

Plurality of
AnalData Stores

2816

Adaptive
Comparator

2806

Storage

2808

HTML

2810

XML

2812

EMAIL

2814

FAX

2818

DBMS

2820

File

2822

Encoded

2824

Compression

2826

2828

2829

2830

2831

2832

2833

2834

2835

2836

2837

2838

2839

2840

2841

2842

2843

2844

2845

2846

2847

2848

2849

2850

2851

2852

2853

2854

2855

2856

2857

2858

2859

2860

2861

2862

2863

2864

2865

2866

2867

2868

2869

2870

2871

2872

2873

2874

2875

2876

2877

2878

2879

2880

2881

2882

2883

2884

2885

2886

2887

2888

2889

2890

2891

2892

2893

2894

2895

2896

2897

2898

2899

2900

2901

2902

2903

2904

2905

2906

2907

2908

2909

2910

2911

2912

2913

2914

2915

2916

2917

2918

2919

2920

2921

2922

2923

2924

2925

2926

2927

2928

2929

2930

2931

2932

2933

2934

2935

2936

2937

2938

2939

2940

2941

2942

2943

2944

2945

2946

2947

2948

2949

2950

2951

2952

2953

2954

2955

2956

2957

2958

2959

2960

2961

2962

2963

2964

2965

2966

2967

2968

2969

2970

2971

2972

2973

2974

2975

2976

2977

2978

2979

2980

2981

2982

2983

2984

2985

2986

2987

2988

2989

2990

2991

2992

2993

2994

2995

2996

2997

2998

2999

29910

29911

29912

29913

29914

29915

29916

29917

29918

29919

29920

29921

29922

29923

29924

29925

29926

29927

29928

29929

29930

29931

29932

29933

29934

29935

29936

29937

29938

29939

29940

29941

29942

29943

29944

29945

29946

29947

29948

29949

29950

29951

29952

29953

29954

29955

29956

29957

29958

29959

29960

29961

29962

29963

29964

29965

29966

29967

29968

29969

29970

29971

29972

29973

29974

29975

29976

29977

29978

29979

29980

29981

29982

29983

29984

29985

29986

29987

29988

29989

29990

29991

29992

29993

29994

29995

29996

29997

29998

29999

299100

299101

299102

299103

299104

299105

299106

299107

299108

299109

299110

299111

299112

299113

299114

299115

299116

299117

299118

299119

299120

299121

299122

299123

299124

299125

299126

299127

299128

299129

299130

299131

299132

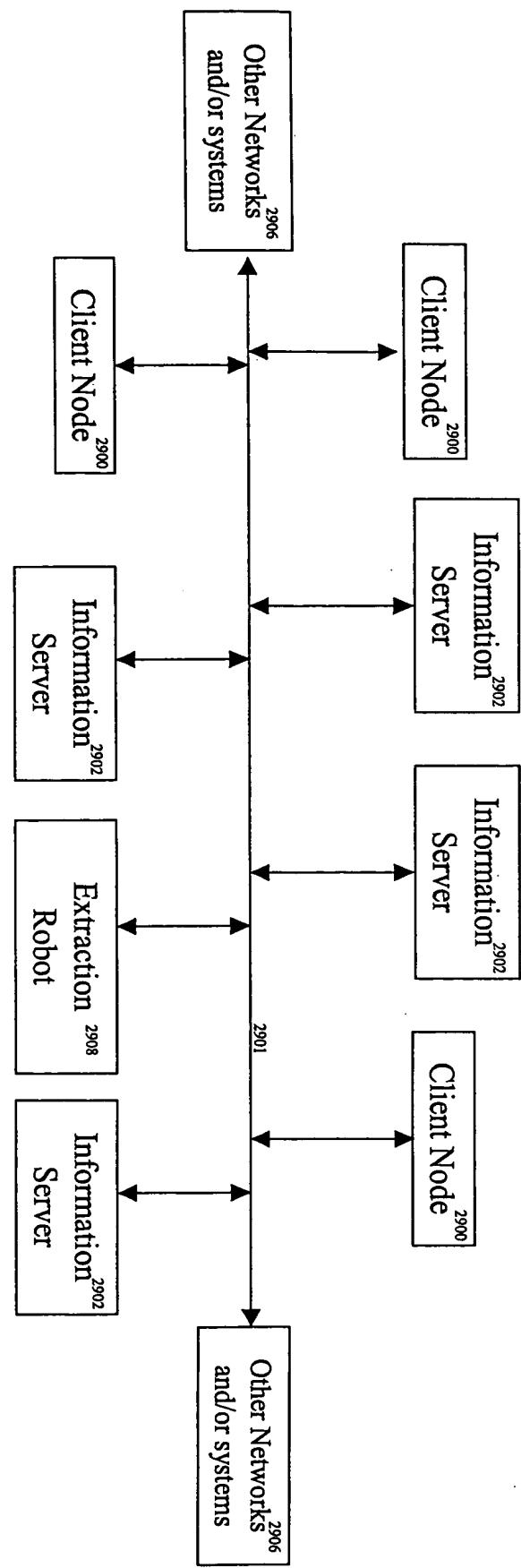
299133

299134

299135

299136

299137



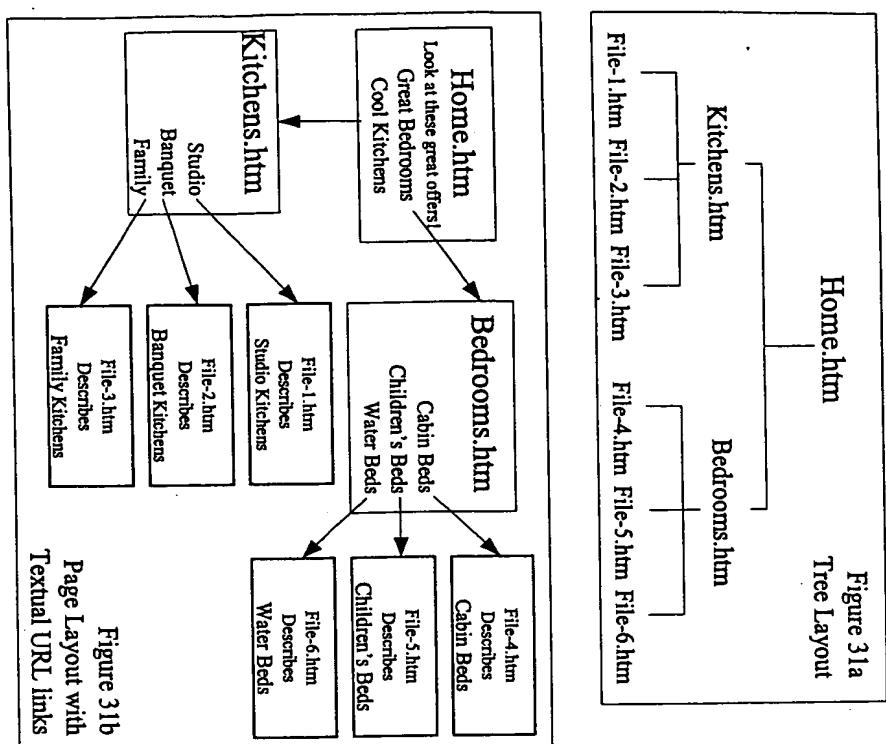
Examples of information servers being accessed from Human operators and mechanized devices.

Figure 29

WWW Page Load, Processing and Display Times	3002
t_{text}^p = Time to obtain, process and display text	
t_{image}^p = Time to obtain, process and display images	
t_{other}^p = Time to obtain, process and display all other items	
$t_{total}^p = t_{text}^p + t_{image}^p + t_{other}^p$ Total time to load, process and display all page items	
Human times to access URL from a Displayed Page	3004
t_h^r = Time to react to and access URL response	
t_h^a = Time for apparatus to respond to URL access	
$t_h^{internal}$	
t_h^{other} = other miscellaneous times	
$t_h^{min} = t_h^r + t_h^a + t_h^{internal}$ Total time to react to and access a URL	
t_h^{max} = infinite.	
Non-Human times to access URL from a Displayed Page	3006
t_{text}^n = Time to obtain, process and display text	
$t_{internal}^n$ = Time for apparatus to respond to URL access (very small)	
t_{other}^n = Time to obtain, process and display all other item (tending to zero)	
$t_{min}^n = t_{text}^n + t_{internal}^n + t_{other}^n$ Total time to react to and access a URL	
t_{max}^n = infinite.	

Figure 30
Timing Definitions

10/517738



Requester ID	Data Item ID	Time Stamp	Type of Access
3118	193.133.51.2	/Home.htm	02/14/00 10:43:15223 Read OK
3120	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:17554 Read OK
3122	193.133.51.2	/Home/Bedrooms/File-4.htm	02/14/00 10:43:20332 Read OK
3124	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:21555 Read OK
3126	193.133.51.2	/Home/Bedrooms/File-5.htm	02/14/00 10:43:24676 Read OK
3128	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:26009 Read OK
3130	193.133.51.2	/Home/Bedrooms/File-6.htm	02/14/00 10:43:29876 Read OK
3132	193.133.51.2	/Home.htm	02/14/00 10:43:31000 Read OK
3134	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:33442 Read OK
3136	193.133.51.2	/Home/Kitchens/File-1.htm	02/14/00 10:43:35998 Read OK
3138	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:37009 Read OK
3140	193.133.51.2	/Home/Kitchens/File-2.htm	02/14/00 10:43:40030 Read OK
3142	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:41993 Read OK
3144	193.133.51.2	/Home/Kitchens/File-3.htm	02/14/00 10:43:45554 Read OK

Example of server access by a <i>human</i> operator				
Requester ID	Data Item ID	Time Stamp	Type of Access	
3150	193.133.51.2	/Home.htm	02/14/00 10:43:15000	Read OK
3152	193.133.51.2	/Home/Bedrooms.htm	02/14/00 10:43:15550	Read OK
3154	193.133.51.2	/Home/Bedrooms/File-4.htm	02/14/00 10:43:15553	Read OK
3156	193.133.51.2	/Home/Bedrooms/File-5.htm	02/14/00 10:43:16000	Read OK
3158	193.133.51.2	/Home/Kitchens.htm	02/14/00 10:43:16005	Read OK
3160	193.133.51.2	/Home/Kitchens/File-2.htm	02/14/00 10:43:16030	Read OK
3162	193.133.51.2	/Home/Kitchens/File-3.htm	02/14/00 10:43:16040	Read OK
3164	193.133.51.2	/Home/Kitchens/File-4.htm		
3166	193.133.51.2	/Home/Bedrooms/File-6.htm		
3168	193.133.51.2	/Home/Bedrooms/File-6.htm		
3170	193.133.51.2	/Home/Kitchens/File-1.htm		
3172	193.133.51.2	/Home/Kitchens/File-2.htm		
3174	193.133.51.2	/Home/Kitchens/File-3.htm		

Example of server access by a *non-human* operator such as an Extraction Robot

Figure 31c

Page hierarchy with textual URL's
Figure 31

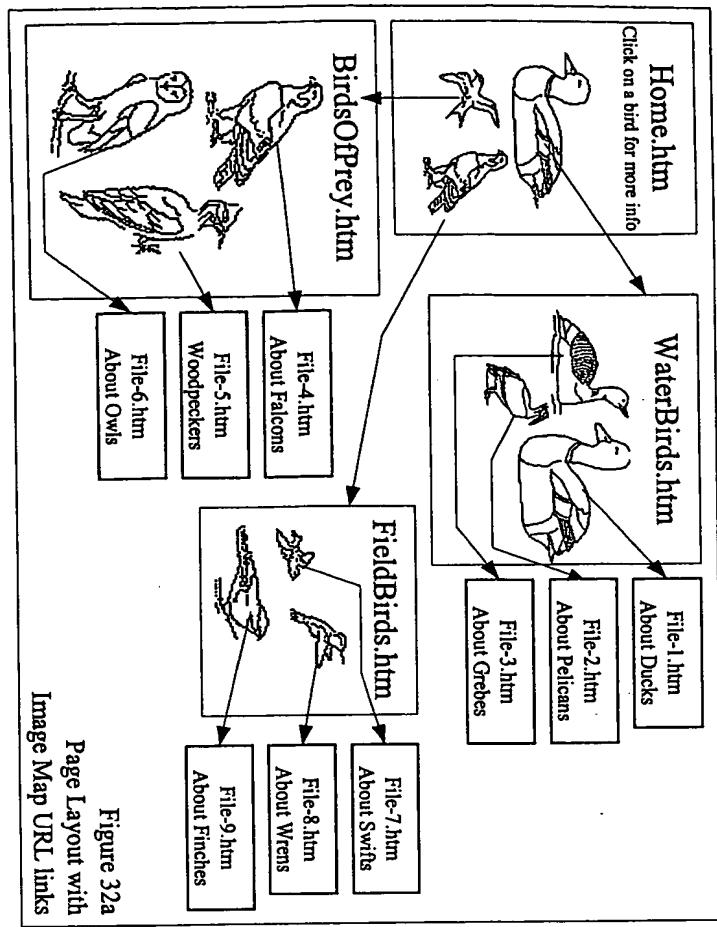


Figure 32
Page Hierarchy with Image Maps

Reference times to access a URL

3300 $t_{\text{ref}}^{\text{response}}$ Time to react to and access URL

3302 $t_{\text{internal}}^{\text{ref}}$ Time for apparatus to respond to URL access

3304 $t_{\text{other}}^{\text{ref}}$ other miscellaneous times

3306 $t_{\text{min}}^{\text{ref}} = t_{\text{response}}^{\text{h}} + t_{\text{internal}}^{\text{h}} + t_{\text{other}}^{\text{h}}$ Total time to react to and access a URL

3308 $t_{\text{max}}^{\text{ref}}$ infinite.

Figure 33a

Timing Definitions

3310 $\Delta \text{hit}_n = \text{hit}_{n+1} - \text{hit}_n$ Time difference between 2 hits

3312 $t_{\text{hit_av}} = \sum_n^n \Delta \text{hit}_n / n$ Average time for hits n_0 to n

3314 $t_{\text{hit_min}}^{\text{no} \rightarrow n}$ Minimum hit time for hits n_0 to n

3316 $t_{\text{hit_max}}^{\text{no} \rightarrow n}$ Maximum hit time for hits n_0 to n

hits n_0 is the first hit in a sequence, e.g. B320

hit n is the last hit in a sequence, e.g. B344

Figure 33b

Human Signature Definitions

3330 $t_{\text{min}}^{\text{h_sig}} = t_{\text{hit_ref}}^{\text{h}}$ Difference between hit time and reference minimum value

3332 $t_{\text{max}}^{\text{h_sig}} = t_{\text{hit_ref}}^{\text{h}}$ Difference between hit time and reference maximum value

3334 $t_{\text{av}}^{\text{h_sig}} = t_{\text{hit_ref}}^{\text{h}}$ Difference between hit time and reference average value

3336 $t_{\text{delta}}^{\text{h_sig}} = (t_{\text{min}}^{\text{h_sig}} + t_{\text{av}}^{\text{h_sig}} + t_{\text{max}}^{\text{h_sig}}) / 3$ Average for all human signature values

Figure 33c

Figure 33
Signature Calculations

Signature Proximity Terms

3400	$t_{\min}^{\text{prox}} = t_{\text{min}}^{\text{h_sig}} - t_{\text{min}}^{\text{r_ref}}$	Difference between human minimum signature and robot reference minimum values
3402	$t_{\text{av}}^{\text{prox}} = t_{\text{av}}^{\text{h_sig}} - t_{\text{av}}^{\text{r_ref}}$	Difference between human average signature and robot reference average values
3404	$t_{\max}^{\text{prox}} = t_{\max}^{\text{h_sig}} - t_{\max}^{\text{r_ref}}$	Difference between human maximum signature and robot reference maximum values
3406	$t_{\min}^{\text{prob}} = t_{\min}^{\text{prox}} \rightarrow t_{\min}^{\text{r_ref}}$	higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.
3408	$t_{\text{av}}^{\text{prob}} = t_{\text{av}}^{\text{prox}} \rightarrow t_{\text{av}}^{\text{r_ref}}$	higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.
3410	$t_{\max}^{\text{prob}} = t_{\max}^{\text{prox}} \rightarrow t_{\max}^{\text{r_ref}}$	higher probability of robot hit. Decreasing positive values and increasing negative values indicate higher probability.

$t_{\text{h_ref}}$ denotes a robot (ie non-human) *hit reference* term
 $t_{\text{h_ref}}$ denotes a human *hit reference* term

Figure 34a

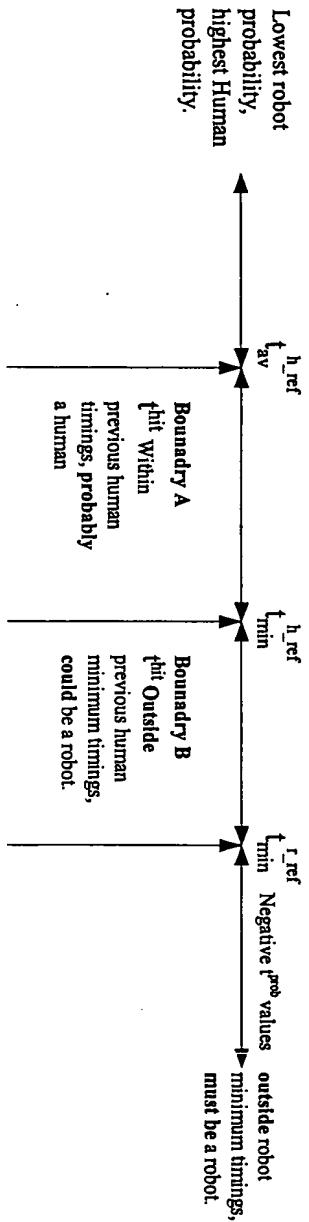


Figure 34b

Figure 34

Signature Proximity Terms

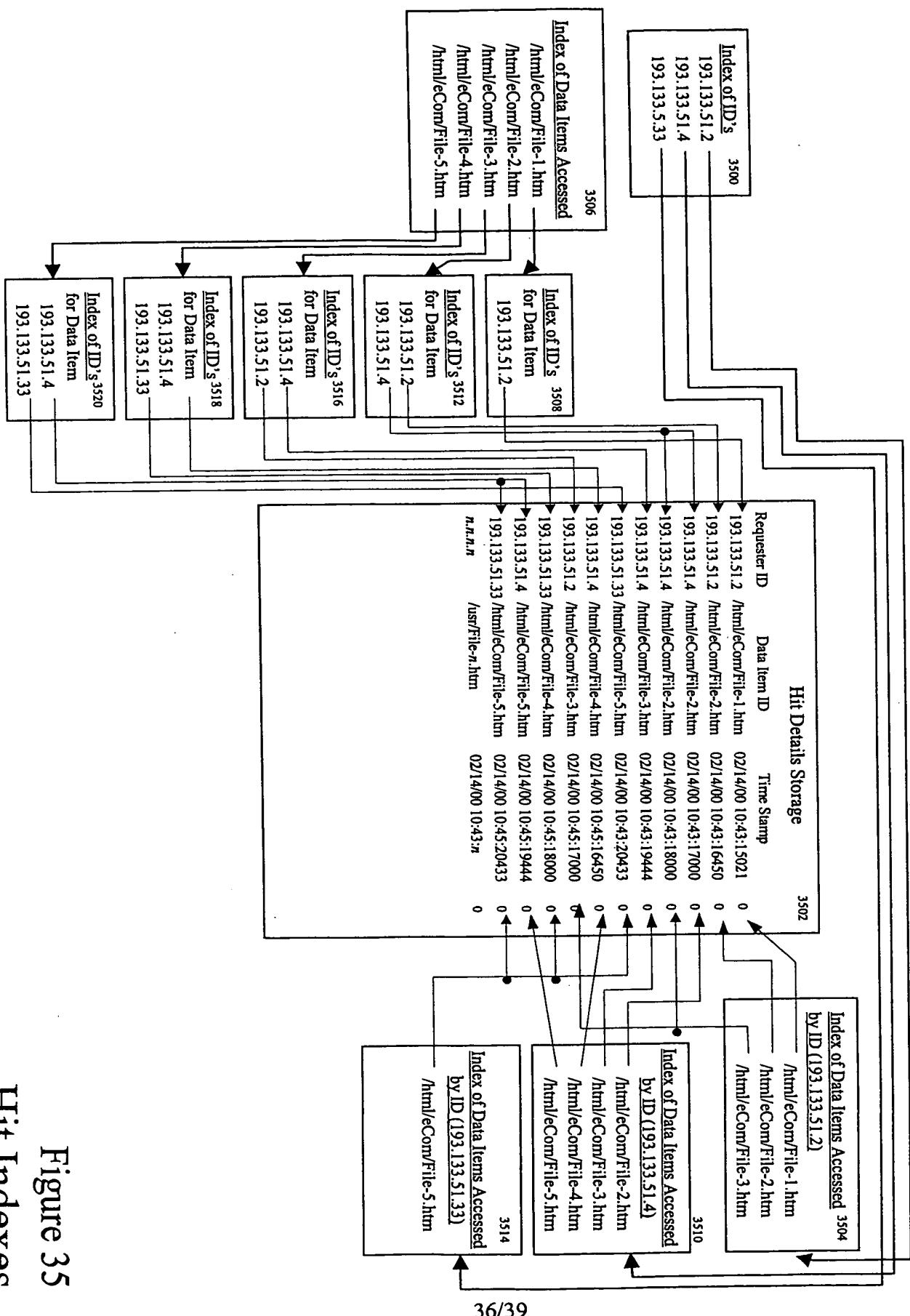
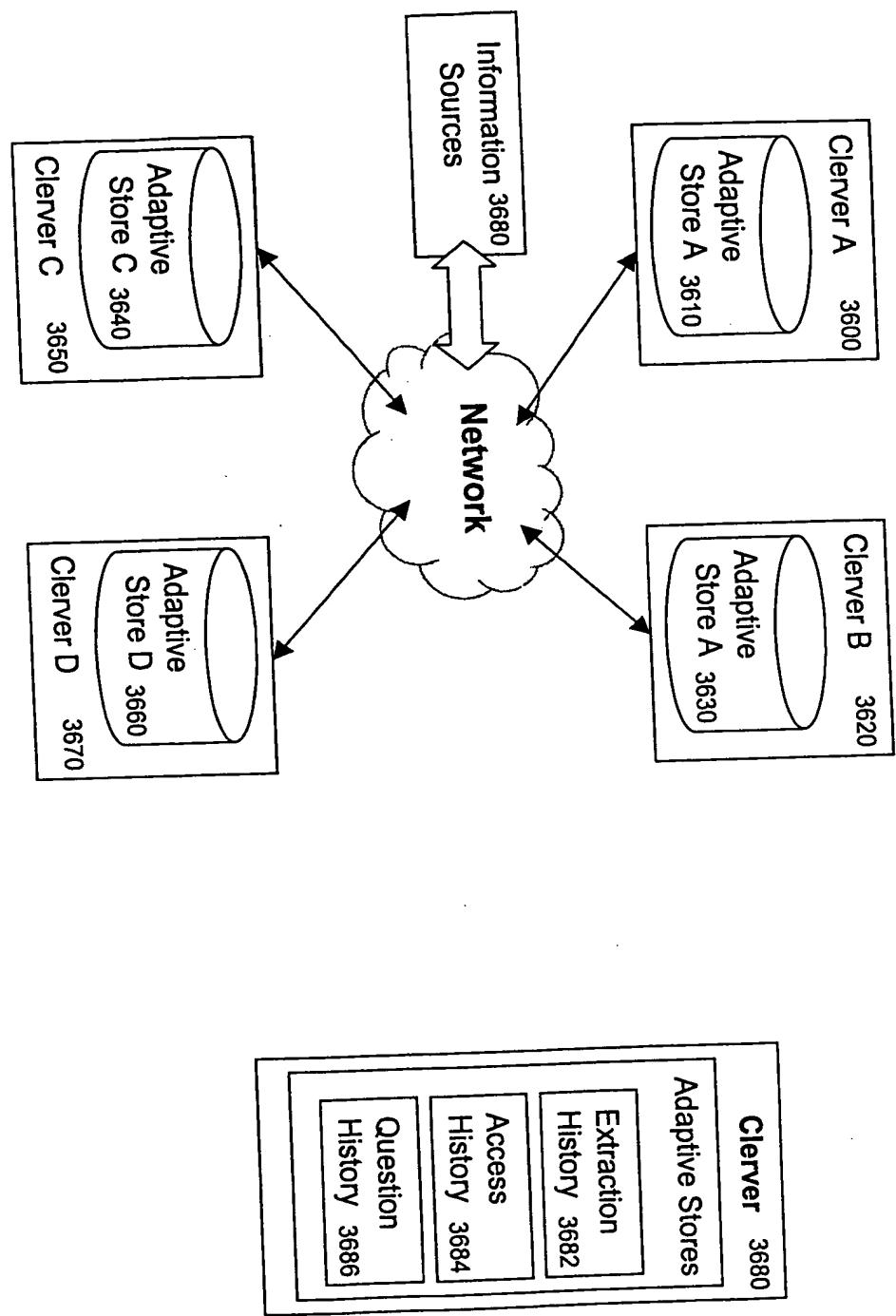
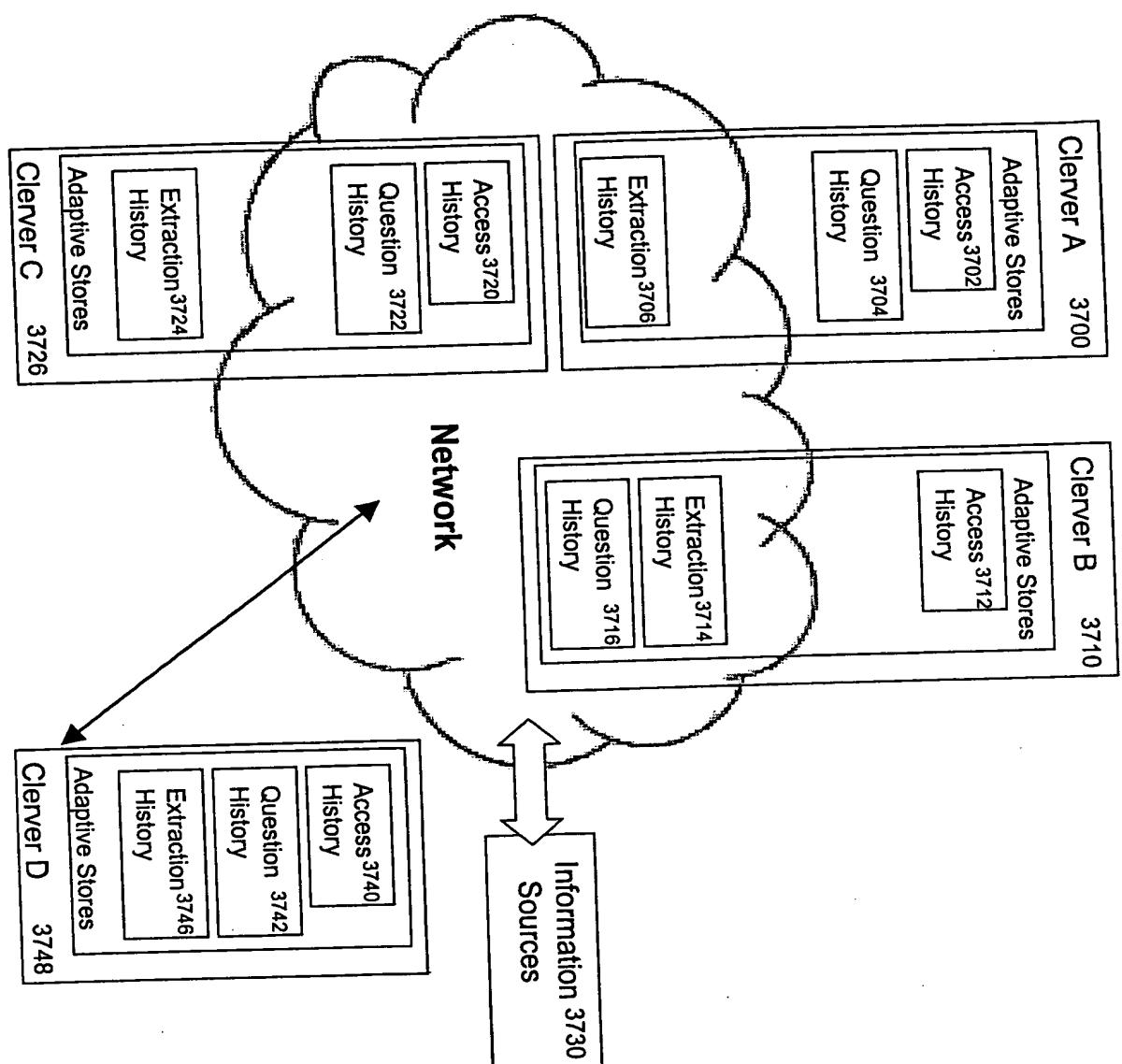


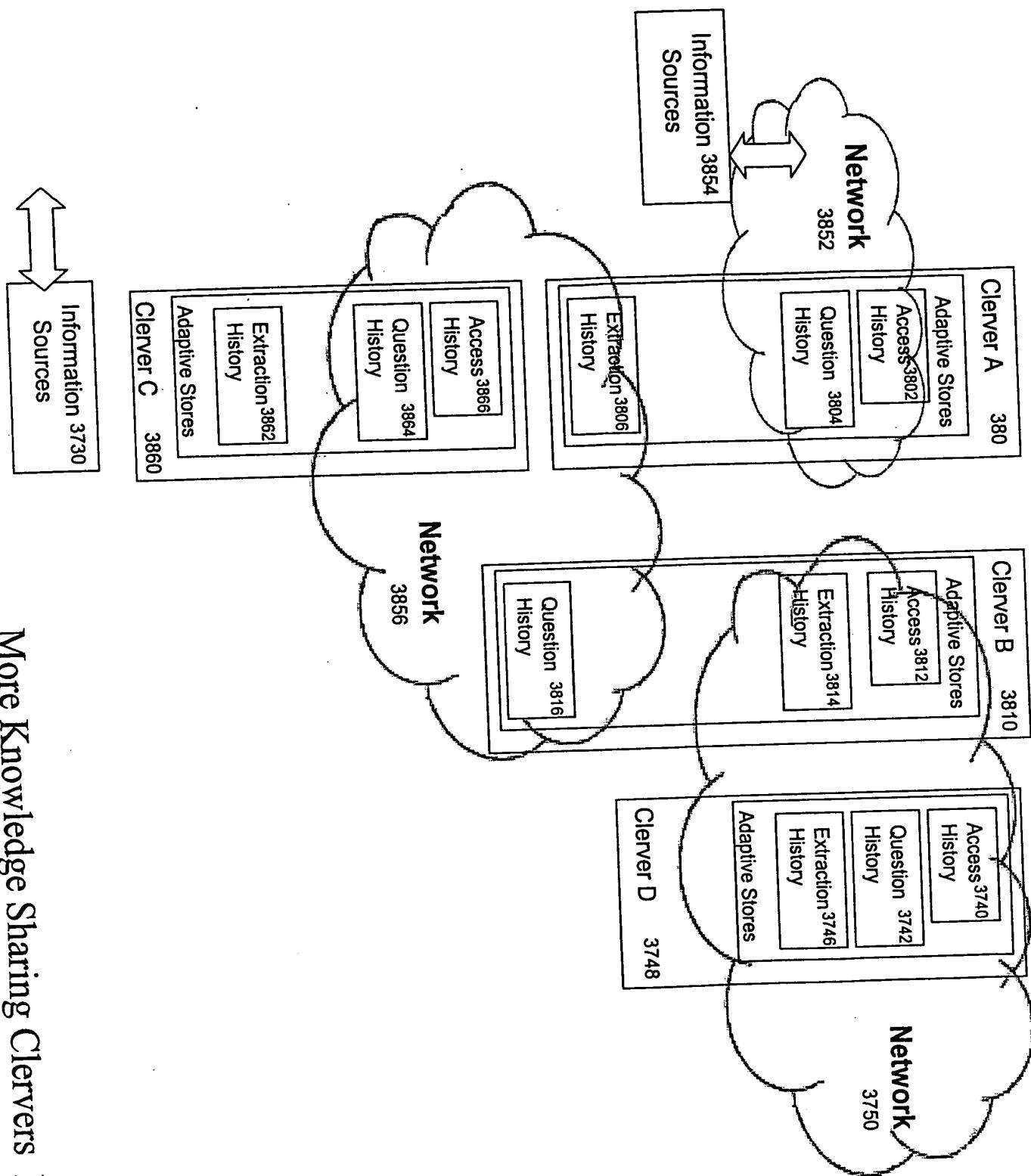
Figure 35
Hit Indexes



Basic Cleverer
Figure 36



Knowledge Sharing Clevers Figure 37



More Knowledge Sharing Clevers Figure 38

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.